BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

REMARKS

The present invention provides RNase P polypeptides and methods for identifying antibiotics using these polypeptides.

Claims 1, 2, 8-11, and 13-40 are pending in this case. Claims 22-28 and 30-40 are withdrawn from consideration by the Examiner. Claims 2, 16, and 20 are rejected under 35 U.S.C. § 112, second paragraph. Claims 1, 2, 8-11, 13-21, and 29 are rejected under 35 U.S.C. § 112, first paragraph. Claim 1 is rejected under 35 U.S.C. § 102, and claims 1, 2, 8-11, 13-21, and 29 are rejected under 35 U.S.C. § 103. Each of these rejections is addressed below in the order that it appears in the Office Action.

Support for the Amendments

Claim 1 has been amended to incorporate the definition of an RNase P consensus sequence found on page 7, lines 1 to 6 of the specification. Claim 8 has been amended to clarify that the agent identified has antibacterial activity. Support for this amendment is found on page 23, lines 25 to 27 of the specification.

Objection to the Drawings and the Specification

The Examiner kindly pointed out an incorrect reference to Figure 2 in the specification. This reference has been corrected and this objection can now be withdrawn. The Examiner also objects to Figure 1 because the residues highlighted in black cannot be seen. The attached substitute drawings have been amended as requested by the Examiner and this objection can now be withdrawn.

Rejection of claims 2, 16, and 20 under 35 U.S.C. § 112, second paragraph

Claims 2, 16, and 20 were rejected under 35 U.S.C. § 112, second paragraph, for reciting non-elected embodiments of the invention. Applicants again respectfully assert that claims 2, 16, and 20 need not be limited to the elected species at this time. Such an amendment would potentially be appropriate when no generic claim is allowed. In the event that a generic claim is allowed, applicants assert that claims to the remaining species, which are written in dependent form, or which otherwise include all the limitations of the allowed generic claim should be considered as provided by 37 CFR § 1.141 and MPEP § 809.02(a). Applicants note that the Examiner has indicated that this rejection will be maintained until a generic claim is allowed or the claims are restricted to the elected species.

Rejection of claims 17 and 21 under 35 U.S.C. § 112, first paragraph

Claims 17 and 21 are rejected, under 35 U.S.C. § 112, first paragraph, for lack of enablement. The Examiner maintains this rejection based on the assertion that that the specification teaches nothing about the enzymatic activity of an *E. coli* or *B. subtilis* RNase P holoenzyme. Applicants respectfully disagree.

The standard for enablement in the biotechnology arts has been set forth in *In re Wands* (858 F.2d 731, 8 U.S.P.Q.2d 1400 (Fed. Cir. 1988)). *Wands* holds that an invention is enabled so long as the teaching of the specification provides the invention without undue experimentation. The fact that experimentation may be complex does not necessarily make it undue, if the art typically engages in such experimentation. *In re Certain Limited-Charge Cell Microcarriers*, 221 U.S.P.Q. 1165, 1174 (Intl. Trade Comm'n 1983), *aff'd sub nom, Massachusetts Institute of Technology v. A.B. Fortia*, 774 F. 2d 1104, 227 U.S.P.Q 428 (Fed. Cir. 1985).

Claims 17 and 21 specify that the enzymatic activity of the *E. coli* or *B. subtilis* RNase P holoenzyme is the hydrolysis of an RNase P substrate. To this end, on page 4, line 25 through page 5, line 10, the specification teaches:

> [b]y "a polypeptide containing RNase P activity" is meant a polypeptide sequence that, when combined with an RNA subunit to form an RNase P holoenzyme, has 20%, 50%, 75%, or even 100% or more, of the enzymatic activity of an *E. coli* or *B. subtilis* RNase P holoenzyme. Preferably, the RNA subunit is from the same species when activity is tested. The enzymatic activity can be assessed, for example, by measuring *hydrolysis of an RNase P substrate.* Standard methods for conducting such hydrolysis assays are described herein and in the literature (see, e.g., Altman and Kirsebom, Ribonuclease P, *The RNA World*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1999; Pascual and Vioque, Proc. Natl. Acad. Sci. 96: 6672, 1999; Geurrier-Takada et al., Cell 35: 849, 1983; Tallsjö and Kirsebom, Nucleic Acids Research 21: 51, 1993; Peck-Miller and Altman, J. Mol. Biol. 221: 1, 1991; Gopalan et al., J. Mol. Biol. 267: 818, 1997; and WO 99/11653).
>
> By "RNase P substrate" is meant a substrate in which hydrolysis by an RNase P holoenzyme requires the presence of the RNase P protein subunit. (emphasis added)

Several of the references cited above describe assays used to measure the enzymatic activity of *E. coli* and *B. subtilis*. For example, the Guerrier-Takada reference describes assay conditions for the hydrolysis of tRNA using both *E. coli* and *B. subtilis* RNase P protein subunits. The Gopalan and Tallsjö references also describe assays for the hydrolysis of tRNA using *E. coli* RNase P protein subunits.

Standard assays for the measurement of hydrolysis of an RNase P substrate by an RNase P holoezyme are further disclosed on pages 19-23 of the specification.

For example, the on page 22, lines 3-21, the specification teaches the following exemplary assays and reaction buffers:

[s]amples of the RNase P holoenzyme and the RNase P substrate are mixed, incubated, and measured for spectrophotometric polarization. When the substrate is cleaved by the RNase P holoenzyme, the 10-nucleotide 5'- leader sequence is released, which leads to a substantial change in the fluorescence polarization in the sample. (Campbell, I.D. & Dwed., R.A. pp. 91-125 The Benjamin/Cummings Publishing Company, Menlo Park, CA (1984); Lakowicz, J.R., Plenum Press, NY (1983)).

The preferred reaction buffer contains 50 mM Tris-HCl (pH 7.5), 100 mM ammonium chloride and 10 mM magnesium chloride. Concentrations of 10-100 mM, 25-500 mM and 1-100 mM of the above, respectively, can be substituted, as can other buffering agents such as MOPS or HEPES, or other monovalent cations, such as sodium or potassium. When the assay is run in either 96 or 384-well polystyrene or polypropylene assay plates, there is a very significant decrease in the fluorescence intensity and polarization of the annealed substrate over time in the absence of enzyme. Various conditions have been tested to prevent the loss of signal with time. The preferred conditions include addition of 10-40 μg/ml carbonic anhydrase and 10-100 μg/ml polyC to the buffer. Other materials, such as, 0.5-5% glycerol, 10-100 μg/ml hen egg lysozyme, 10-50 μg/mL tRNA, 1-10 mM DTT, or 2-10 mM DTT can also be added to the buffer to prevent some loss of signal.

In addition, the previously submitted Declaration of Dr. Gopalan states that a skilled artisan can easily measure the rate of hydrolysis of an RNase P substrate by an RNase P holoezyme of interest and determine whether that rate is at least 20% of the rate of hydrolysis of the same RNase P substrate by the same concentration of an *E. coli* or *B. subtilis* RNase P holoenzyme under the same conditions, using standard assays such as those described in the specification. Exhibit 1, submitted with the previously submitted Declaration of Dr. Gopalan, provides support for this assertion by showing the results of an exemplary assay, using conditions provided in the specification. In this example, the activity of an RNase P polypeptide of interest (*N. gonorrhoeae;* SEQ ID NO: 27 of the present application) was compared

to that of *E. coli* RNase P to determine whether the activity of *N. gonorrhoeae* RNase P is at least 20% of the activity of *E. coli* RNase P. The reaction mixture included 1 nM RNase P RNA subunit, 1-5 nM RNase P protein subunit, 40 nM pre-tRNA substrate, 50 mM Tris-HCl, pH 7.5, 100 mM NH_4Cl, and 10 mM $MgCl_2$ at room temperature (22 °C). The reaction was allowed to proceed for 5-60 minutes.

In the present invention, the enzymatic activity of a polypeptide containing RNase P activity identified in the present invention is compared to the enzymatic activity of an *E. coli* or *B. subtilis* RNase P enzyme. Quantitation of the resulting signal or bands on the autoradiogram is required for such a comparison. Quantitaion of the resulting signal or bands is a standard art-known procedure which can be accomplished, for example, by direct exposure of the gel to a phosphorous screen in a phosphorimager that allows for quantitation of the radioactive signal itself (see for example Gopalan et al. cited above). Such quantitation methods provide relative and not absolute values depending on several variables including the optimization of the assay conditions used and exposure time of the gel to the autoradiogram or of the gel to the phosphorimager. Again, as stated in the Declaration of Dr. Gopalan, the skilled artisan can compare the relative enzymatic activity using the same substrate and the same concentration of holoenzyme, under the same conditions. Since the present invention requires the determination of *relative enzymatic activity*, the exact assay conditions need not be replicated precisely but can be optimized, using standard art-known methods, such as those provided in the references and citations above, for each assay. Optimization of assay conditions is a standard procedure in the art and does not constitute undue experimentation.

In summary, using the teachings of the specification and routine methods for quantitation known in the art, a skilled artisan could readily assay additional

polypeptides having RNase P consensus sequence for RNase P activity and determine if the RNase P polypeptide of interest has at least 20% of the enzymatic activity of an *E. coli* or *B. subtilis* RNase P polypeptide. The standard set forth in *Wands* that a claimed invention must be enabled so that any person skilled in the art can make and use it without undue experimentation has been met in the present case. Accordingly, this rejection can now be withdrawn.

Rejection of claims 1, 2, 8-11, 13-21, and 29 under 35 U.S.C. § 112, first paragraph

Claims 1, 2, 8-11, and 13-21 are rejected under 35 U.S.C. § 112, first paragraph, for failure to convey possession of the claimed invention. The Examiner states that in order to adequately describe the instant claims, one would have to provide the sequences of all 59 of the excluded enzymes. Applicants respectfully point out that the entire sequence of each of the 59 excluded enzymes is included by reference in the specification on page 7, line 17-22.

> The RNase P sequences claimed as part of the present invention specifically exclude those sequences in the RNase P database (James W. Brown, The Ribonuclease P Database, Nucleic Acids Research 27(1):314 (1999)) as posted on the internet on March 1, 2000. Also excluded are the RNase P polypeptide and nucleic acids described by nucleic acid or amino acid sequence in EP 0811 688 A2 (*Staphylococcus aureus*) and WO 99/11653 (*S. pneumoniae*).

As stated in the declaration of Dr. Vicki Healy, submitted on October 11, 2002, the protein sequences for each of the RNase P protein subunits listed in claim 1, except for *Staphylococcus aureus* and *S. pneumoniae*, were available in the RNase P Database of James W. Brown on March 1, 2000. The sequence for the RNase P protein subunits of *Staphylococcus aureus* and *S. pneumoniae* are provided by reference in the patent applications also cited above. For clarity, applicants have

also provided an enlarged version of Figure 1 (Exhibit A), which shows the entire sequence of each of the RNase P protein subunits listed. Applicants assert that claim 1 is adequately described because the complete sequences of the excluded RNase P protein subunits are provided in the specification.

The Examiner also states that there needs to be some rule for determining which polypeptides are RNase P enzymes in the instant claims. Applicants respectfully assert that such a rule is clearly found in the specification, for example, on pages 10-11.

> To identify which sequences were genuine RNase P protein subunits, we determined whether the sequences also contained an RNase P consensus sequence, which we defined as a sequence that, upon alignment with known RNase P sequences using the ClustalW program, conserves at least nine of the following twenty amino acids in the *E. coli* RNase P protein sequence: R11, L12, F18, R46, G48, V51, K53, K54, A59, V60, R62, N63, K66, R67, R70, L80, D84, V86, L101, and L105 (page 10, lines 19-25).
>
> Based upon these known sequences, we determined that a polypeptide identified by our above-described RNase P BLAST search contained an RNase consensus sequence and was a genuine RNase P protein subunit if it contained at least nine of the above-described twenty amino acids (page 11, lines 12-15).

However, in order to expedite the prosecution of this application, applicants have amended claim 1 to include the rule for determining which polypeptides are genuine RNase P subunits.

As described on page 11, lines 18 to 25, the preferred subset of 9 out of the 20 amino acids is,

> preferred because it has been identified as playing a significant role in RNase P function through mutation studies (Gopalan et al., J. Mol. Biol. 267:818 1997) and the determination of the RNase P three dimensional

structure (Stams et al., Science 280: 752, 1998). The three dimensional structure reveals that all of the residues that make up the above-described nine amino acid subset are proximal to each other in the tertiary structure of the protein, despite the distance between some of the residues in the primary structure.

The nine residues specified as a preferred subset of the twenty amino acids on, for example, page 7, lines 9 to 10, of the specification (F18, R46, K53, A59, R62, N63, K66, R67, and R70) rank among the highest in terms of identity and specificity. Using the tertiary structures of RNase P proteins of *Bacillus subtilis* and *Staphylococcus aureus*, Jovanovic et al. (Nucleic Acids Research 30: 5065-5073, 2002; attached as Exhibit B) determined that the conserved amino acids of RNase P can be grouped into two specific conserved regions, the helix 2 region and the large central cleft formed by packing of α1 against the β-sheet. Figure 2A and 2B of Jovanovic et al., and the enlarged version of the figures (attached as Exhibit C) show each of the twenty conserved residues and their location in the conserved α-helix and the large central cleft. Although these residues do not appear sequentially when the protein sequence is presented in a linear fashion, spatially they are present on the same two surfaces, thereby demonstrating conservation of both relative sequence position and structural location. The high degree of conservation of these amino acids both in terms of sequence and in terms of spatial location underscores the importance of these nine amino acids in the general function of an RNase P polypeptide and provides evidence of the significance of the specified 9 and 20 amino acid residues. As described above, both the 20 amino acid residues and the subset of 9 amino acid residues are conserved throughout all of the bacterial RNase P proteins shown in Figure 1 both in terms of sequence and structure, supporting the factual basis for the use of these conserved amino acids to identify additional RNase P protein subunits.

The Examiner also rejects claims 1, 2, 8-11, 13-21, and 29 under 35 U.S.C. § 112 for failure to comply with the written description requirement by stating that, although the specification makes the presumption that an enzyme with similarity to

a known RNAse P enzyme will also have activity, this presumption was never tested. Applicants respectfully disagree and assert that there are indeed examples, found in the specification and in the previously submitted Declaration of Dr. Gopalan, of the identification of the catalytic activity of an RNase P protein using the consensus sequence and that such examples are sufficient to satisfy the written description requirement under 35 U.S.C. § 112.

The written description requirement, as set forth in 35 U.S.C. § 112, first paragraph, requires that the "specification shall contain a written description of the invention." The M.P.E.P. § 2163 states:

> The written description requirement has several policy objectives. "[T]he 'essential goal' of the description of the invention requirement is to clearly convey the information that an applicant has invented the subject matter which is claimed." *In re Barker*, 559 F.2d 588, 592 n.4, 194 USPQ 470, 473 n.4 (CCPA 1977). Another objective is to put the public in possession of what the applicant claims as the invention. See *Regents of the University of California v. Eli Lilly*, 119 F.3d 1559, 1566, 43 USPQ2d 1398, 1404 (Fed. Cir. 1997), *cert. denied*, 523 U.S. 1089 (1998).
>
> An applicant shows possession of the claimed invention by describing the claimed invention with all of its limitations using such descriptive means as words, structures, figures, diagrams, and formulas that fully set forth the claimed invention. *Lockwood v. American Airlines, Inc.*, 107 F.3d 1565, 1572, 41 USPQ2d 1961, 1966 (Fed. Cir. 1997). Possession may be shown in a variety of ways including *description of an actual reduction to practice*, or by showing that the invention was "ready for patenting" such as by the disclosure of drawings or structural chemical formulas that show that the invention was complete, or by describing distinguishing identifying characteristics sufficient to show that the applicant was in possession of the claimed invention. See, e.g., *Pfaff v. Wells Elecs., Inc.*, 525 U.S. 55, 68, 119 S.Ct. 304, 312, 48 USPQ2d 1641, 1647 (1998); *Eli Lilly*, 119 F.3d at 1568, 43 USPQ2d at 1406; *Amgen, Inc. v. Chugai Pharmaceutical*,

> 927 F.2d 1200, 1206, 18 USPQ2d 1016, 1021 (Fed. Cir. 1991) (one must define a compound by "whatever characteristics sufficiently distinguish it"). (M.P.E.P. § 2163; emphasis added.)

Applicants' specification describes both the distinguishing characteristics of the invention, including the rule for identifying an RNase P holoenzyme using the RNase P consensus sequence described above, as well as examples of an actual reduction to practice showing the application of this rule to identify an RNase P protein subunit having enzymatic activity.

> Further support for the RNase P identification is as follows. Using the above-described BLAST search and consensus sequence determination, we independently identified the sequence for an RNase P protein subunit from the genomic database of *Staphylococcus aureus* (*S. aureus*). Although this sequence had been previously identified as an RNase P protein subunit and its RNase P activity had been confirmed by assay (EPA 0 811 688 A2), our independent discovery of this RNase P sequence provides proof of principle that our method of searching for RNase P protein subunits predictably identifies polypeptides that have RNase P activity. (page 12, line 25 to page 13, line 4)

Clearly, the citation above, taken from the specification, describes a test of the hypothesis that an enzyme identified using the rule for identifying a genuine RNAse P polypeptide, as cited above, will have activity.

In addition, the previously submitted Declaration of Dr. Gopalan, also described above, shows evidence that two RNase P polypeptides (*N. gonorrhoeae* and *Porphyromonas gingivalis*, SEQ ID NOS: 27 and 31, respectively), identified by the methods provided in the specification, have RNase P enzymatic activity. These three examples support the proposition that an RNase P holoenzyme

reconstituted using a polypeptide identified by the methods described in the present application will have enzymatic activity.

According to the "Guidelines for the Examination of Patent Applications Under the 35 U.S.C. § 112, para. 1, 'Written Description' Requirement" set forth in § 2163 of the MPEP,

> [a] "representative number of species" means that the species which are adequately described are representative of the entire genus. Thus, when there is substantial variation within the genus, one must describe a sufficient variety of species to reflect the variation within the genus. On the other hand, there may be situations where one species adequately supports a genus.

Given that the methods for identifying an RNase P protein subunit and determining if the protein has RNase P activity are clearly presented in the specification and have very little, if any variation in their application to the various species claimed, the representative number of species showing RNase P activity for the identified RNase P protein subunits meets the guidelines for the written description requirement presented above. Applicants assert that for the reasons outlined above, the specification demonstrates that applicants were in possession of the invention at the time the application was filed. Accordingly, this rejection should be withdrawn.

Rejection of claims 1, 2, 8-11, 13-21 and 29 under 35 U.S.C. § 112, first paragraph

Claim 1, 2, 8-11, 13-21, and 29 were rejected, under 35 U.S.C. § 112, first paragraph, for lack of enablement. The Examiner maintains this rejection because, "there is no rule in any of the claims as to exactly how many of the residues must be identical in order for the polypeptide to be an RNase P, nor is there any disclosure that the RNase P polypeptides identified in the invention as SEQ ID NOs: 20-38 have RNase P activity." Applicants respectfully disagree.

Applicants have amended claim 1 to specifically recite an isolated polypeptide comprising "at least nine of the following specified 20 amino acid residues in the *E. coli* RNase P protein subunit: R11, L12, F18, R46, G48, V51, K53, K54, A59, V60, R62, N63, K66, R67, R70, L80, D84, V86, L101, and L105." This definition of the term "RNase P consensus sequence" is found on page 7, lines 1 to 6 of the specification.

> By "an RNase P consensus sequence" is meant a sequence which, when aligned to the *E. coli* RNase P sequence using the ClustalW program and performing a comparison of the specified amino acid sequences, shows conservation of at least nine of the following specified 20 amino acid residues in the *E. coli* RNase P protein subunit: R11, L12, F18, R46, G48, V51, K53, K54, A59, V60, R62, N63, K66, R67, R70, L80, D84, V86, L101, and L105.

In view of this amendment, this rejection may now be withdrawn.

The Examiner also suggests that the claims are not enabled because there is apparently no disclosure that SEQ ID NOs: 20-38 have RNase P activity.

As stated above, the standard for enablement in the biotechnology arts has been set forth in *In re Wands* (858 F.2d 731, 8 U.S.P.Q.2d 1400 (Fed. Cir. 1988)). *Wands* holds that an invention is enabled so long as the teaching of the specification provides the invention without undue experimentation. *Wands* states that:

> the test [for determining whether experimentation is undue] is not merely quantitative, since a considerable amount of experimentation is permissible, <u>if it is merely routine or if the specification in question provides a reasonable amount of guidance with respect to the direction in which the experimentation should proceed</u> (emphasis added).

Furthermore, the Federal Circuit has long held that it is not necessary for all possible embodiments of a claim to be operative in order for that claim to be enabled. *See Atlas Powder Co. v. E.I. du Pont de Nemours & Co.*, 750 F.2d 1569, 224 U.S.P.Q. (Fed. Cir. 1984). The proper test of enablement is "whether one reasonably skilled in the art could make or use the invention from the disclosures in the patent coupled with the information known in the art without undue experimentation." *Hybritech, Inc. v. Monoclonal Antibodies, Inc.* 802 F.2d. 1318 (Fed. Cir. 1985).

Applying these standards to the present case, it is clear that applicants' specification satisfies the proper test of enablement as outlined above. As described above with reference to the written description prong of the § 112 rejections, several examples of the enzymatic activity of RNase P protein subunits identified in the present invention have been provided. First, the specification describes the identification of the RNase P protein subunit of *Staphylococcus aureus* (*S. aureus*), whose RNase P activity had been previously confirmed by assay (EPA 0 811 688 A2). The previously submitted Declaration of Dr. Gopalan illustrates the ability of recombinant *N. gonorrhoeae* (SEQ ID NO: 27) and *Porphyromonas gingivalis* (SEQ ID NO: 31) RNase P to cleave a pre-tRNA^{Gln} substrate using the methods similar to those described in the specification. These results provide confirmation that polypeptides identified using the methods of the present invention do, in fact, have RNase P activity.

The Examiner acknowledges the previously submitted Declaration of Dr. Gopalan but asks that applicants' provide further proof that there is only one RNase P protein subunit for each bacterial species.

The attached second Declaration of Dr. Gopalan submits data from the Clusters of Orthologous Groups of Proteins (COG) database (www.ncbi.nlm.nih.gov/cog) that provides further proof that there is only one

RNase P protein subunit for each bacterial species. The COG database is a government-supported database that provides a phylogenetic classification of the proteins encoded in complete genomes including bacterial, archael, and eukaryotic genomes. Each cluster group, or COG, is assigned a number and consists of an individual protein from at least 3 lineages. The RNase P polypeptide subunit has been assigned COG594. A search of the 131 genomes of the COG database for COG594 was performed using the Comparative Genomics web tool found at the web site for the Virtual Institute for Microbial Stress and Survival (VIMSS) supported by the U.S. Department of Energy Office of Science (http://vimss.lbl.gov/). This search produced 105 hits (a copy of these results is attached as Exhibit D). Each of the bacteria listed in the 105 hits shows only one complete sequence for the RNase P polypeptide subunit. This tool searches only genomes that have been completely sequenced and would therefore detect more than one copy of an RNase P protein subunit in any given genome. Included in the 105 hits are the RNase P polypeptide subunits for bacteria listed with multiple accession numbers in claim 1. Examples of these include *Streptomyces coelicolor* (A3(2)) and *Mycobacterium tuberculosis* (H37Rv). The data from the COG database provides additional support for the assertion that there is only one RNase P polypeptide subunit for each bacterial species.

The Examiner also points out that there is no statement in the Declaration of Dr. Gopalan that the *N. gonorrheae* and *Porphyromonas gingivalis* enzymes assays are the same as SEQ ID NOs: 27 and 31. A second Declaration of Dr. Gopalan is attached, attesting to the fact that the enzymes used in the assays described in the previously submitted Declaration, are, in fact, the exact same enzymes as SEQ ID NOs: 27 and 31.

In conclusion, the facts in the present case, when combined with the additional declaration as requested by the Examiner, demonstrate that the specification clearly satisfies 35 U.S.C. § 112.

Rejection of claims 8-11, 13-14 and 18-21 under 35 U.S.C. § 112, first paragraph

Claim 8-11, 13-14, and 18-21 are further rejected under 35 U.S.C. § 112, first paragraph for lack of enablement. The Examiner states that while claim 8 defines what is considered an antibiotic agent, the specification does not teach that such an agent is an antibiotic agent that will inhibit the growth of a microorganism, as is generally recognized in the art. Therefore, the Examiner concludes that the specification does not enable one of ordinary skill in the art to identify an antibiotic agent.

Amended claim 8 is directed to a method of identifying an agent which may be useful as an antibacterial agent and is reproduced below.

> 8. A method of identifying an agent, which may be useful as an antibacterial agent, said method comprising:
> i) obtaining an RNase P holoenzyme comprising the polypeptide of claim 1;
> ii) contacting said holoenzyme with an RNase P substrate in the presence and in the absence of a compound; and
> iii) measuring the enzymatic activity of said holoenzyme;
> wherein a compound is identified as an agent which may be useful as an antibacterial agent if said compound produces a detectable decrease in said RNase P enzymatic activity as compared to activity in the absence of said compound.

The specification describes assays used to screen for compounds that inhibit the activity of the RNase P holoenzymes. RNase P is a key enzyme involved in the biosynthesis of tRNA. RNAse P is required for bacterial cell viability *in vivo*.

According to Stedman's Medical Dictionary, 25th edition, an antibiotic is, "a soluble substance derived from a mold or bacterium that inhibits the growth of other microorganisms." By definition then a compound that inhibits the activity of an enzyme required for cell growth would destroy or inhibit the growth of a microorganism. If an enzyme is required for cell growth, then inhibition of the enzyme would inhibit cell growth, which satisfies the definition of an antibiotic. Clearly, a screen for compounds that inhibit the activity of the RNase P holoenzyme is intended to identify compounds that fall under the art known definition of an antibiotic.

However, to further clarify this point, applicants have amended claim 8 to specify that the agent may be useful as an antibacterial agent. Support for this amendment can be found on page 23, lines 25-27 of the specification, which describes the antibacterial activity of inhibitory compounds identified using the method of claim 8.

> Such inhibitors have the advantage of providing a selective *antibacterial treatment* that reduces the adverse side effects associated with killing nonpathogenic bacteria. (emphasis added)

In view of the amendment to claim 8, and for the reasons outlined above, claims 8-11, 13-14, and 18-21 are clearly enabled and this rejection should be withdrawn.

Rejection of claim 1 under 35 U.S.C. § 102

Claim 1 is rejected, under 35 U.S.C. § 102(b), as being anticipated by Gress (WO 99/11653), Guth (EP 0 811 688), Altman (The RNA World, 2:1155-1184, 1999, and FASEB Journal, 7:7-14, 1993), Frank (*Annu. Rev. Biochem.*, 67:153-180, 1998), Gopalan (*J. Mol. Biol.*, 267:818-829, 1997), Pace (*J. Bacteriol.*, 177:1919-1928, 1995), Pascual (*Proc. Natl. Acad. Sci. USA*, 96:6672-6677, 1999), or Peck-

Miller (*J. Mol. Biol.*, 221:1-5, 1991). The Examiner maintains this rejection because, according to the Examiner, applicants have not shown that the RNase P in the database is that excluded by the instant claim.

These above-cited references focus on the following bacterial RNase P subunits or complexes from the following bacteria: *S. pneumoniae* (Gress), *Staphylococcus aureus* (Guth), *E. coli* (Altman, 1999), *E. coli*, *Baccillus subtilis*, *Proteus mirabilis*, *Streptomyces bikiniensis*, and *Micrococcus luteus* (Altman, 1993), *E. coli*, *Baccillus subtilis* (Frank), *E. coli, Buchnera aphidocola, Coxiella burnstii, Haemophilus influenzae, Proteus mirabilis, Pseudomonas putida, Mycoplasma capricolum, Mycobacterium leprae, Micrococcus luteus, Streptomyces coelicolor*, and *Bacillus subtilis* (Gopalan), *E. coli* and *Baccillus subtilis* (Pace), *E. coli* and *Synechocystis* (Pascual), and *E. coli* (Peck-Miller). As described above and in the previous Declaration of Dr. Gopalan, there is only one RNase P polypeptide in each bacterial species. The different accession numbers for some RNase P polypeptide sequences is due to multiple deposits of the same sequence. The Examiner has requested that additional proof to support this statement.

Applicants submit the attached second Declaration of Dr. Gopalan demonstrating a search of the 131 genomes of the COG database for COG594 (the assigned number for the RNase P protein subunit) using the Comparative Genomics web tool found at the web site for the Virtual Institute for Microbial Stress and Survival (VIMSS) supported by the U.S. Department of Energy Office of Science (http://vimss.lbl.gov/). This tool searches only genomes that have been completely sequence and would therefore detect more than one copy of an RNase P protein subunit in any given genome. This search produced 105 hits (a copy of these results is attached as Exhibit D). Each of the bacteria listed in the 105 hits shows only one complete sequence for the RNase P polypeptide subunit. Included in the 105 hits are the RNase P polypeptide subunits for bacteria listed with multiple accession numbers in claim 1. Examples of these include *Streptomyces*

coelicolor (A3(2)), *B. subtilis*, and *Mycobacterium tuberculosis* (H37Rv). The data from the COG database provides additional support for the assertion that there is only one RNase P polypeptide subunit for each bacterial species.

In addition, claim 1 specifically states, "wherein said polypeptide is not a polypeptide from one of the following organisms...(emphasis added)." While applicants again point out that there is only one RNase P polypeptide subunit for each bacterial species, the claim excludes any RNase P polypeptides from the specified organisms, regardless of the number of such RNase P polypeptides. As a result, any and all RNase P polypeptides from the specified organisms would be excluded as a limitation of claim 1.

In view of these clarifying remarks, it should now be clear that the RNase P polypeptides disclosed in the above-mentioned references do not fall within the scope of claim 1 and this rejection may now be withdrawn.

Rejection of claims 8, 10, 11, 13, and 14 under 35 U.S.C. § 103(a)

Claims 8, 10, 11, 13, and 14 are rejected, under 35 U.S.C. § 103(a), as being unpatentable over Potuschak (*Nucl. Acids Res.* 21:3229-3243, 1993), Mikkelsen (*Proc. Natl. Acad. Sci., USA,* 96:6155-6160, 1999), or Schroeder (*EMBO J.,* 19(1): 1-9, 2000), in view of Spitzfaden (*J. Mol. Biol.* 295:105-115, 2000). The Examiner states that it would have been obvious to identify an antibiotic agent by seeing if it decreased the activity of RNase P on a RNase P substrate in view of the references above.

Claim 8 is directed to a method of identifying an antibiotic agent using an RNAse P holoenzyme that includes a polypeptide of claim 1. Polypeptides of claim 1 have an RNase P consensus sequence, which as outlined above, is defined as, "a sequence which...shows conservation of at least nine of the following specified 20 amino acid residues in the *E. coli* RNase P protein subunit: R11, L12, F18, R46,

G48, V51, K53, K54, A59, V60, R62, N63, K66, R67, R70, L80, D84, V86, L101, and L105." None of the references cited by the examiner describe a polypeptide having an RNase P consensus sequence or the use of such a polypeptide to identify an antibiotic agent.

For the reasons provided below, applicants contend it would not have been *prima facie* obvious to use a bacterial RNase P holoenzyme including an RNase P polypeptide of the invention to identify antibiotic agents, as claimed in claims 8, 10, 11, 13, and 14. To establish a *prima facie* case of obviousness of a claimed invention, all claim limitations must be taught or suggested by the prior art. *In re Royka*, 490 F.2d 981, 180 U.S.P.Q. 580 (C.C.P.A. 1974).

> To establish a prima facie case of obviousness, three basic criteria must be met. First, there must be some suggestion of motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. Second, there must be a reasonable expectation of success. Finally, the prior art reference (or references when combined) must teach or suggest all the claim limitations. The teaching or suggestion to make the claimed combination and the reasonable expectation of success must both be found in the prior art, and not based on applicant's disclosure. *In re Vaeck*, 947 F.2d 488, 20 U.S.P.Q.2d 1438 (Fed. Cir. 1991).

These criteria are not satisfied in the present rejection of claims 8, 10, 11, 13, and 14, for obviousness.

Spitzfaden describes the structure determination of the RNase P protein from *Staphylococcus aureus* by NMR spectroscopy and the localization of the RNA binding site. The RNase P polypeptide subunit of *S. aureus* is specifically excluded by claim 1. Spitzfaden does not describe an RNase P consensus sequence or a polypeptide having such a sequence, nor does Spitzfaden identify the 20 critical

amino acids or the subset of 9 amino acids required by the present claims to identify an RNase P protein subunit. Finally, Spitzfaden does not teach the use of an RNase P holoenzyme having such a polypeptide subunit to identify antibiotic agents.

Potuschak explores the similarities between two mouse ribonucleoproteins, RNase MRP and RNase P using two different antibiotics, puromycin and cycloheximide. Potuschak demonstrates that puromycin can inhibit the activity of both mouse enzymes while cycloheximide has no effect. Potuschak also uses *E. coli* RNase P holoenzyme purified from a crude bacterial extract to demonstrate that *E. coli* RNase P can cleave the substrate for RNase MRP and again demonstrates that the two enzymes are similar in their ability to cleave the same substrate.

The Potuschak reference does not disclose a polypeptide having an RNase P consensus sequence nor does it teach the use of a bacterial RNase P holoenzyme having such a polypeptide subunit to identify antibiotic agents. In addition, claim 1 specifically excludes the *E. coli* RNase P used by Potuschak.

The Schroeder reference is a review that covers the modulation of RNA function by antibiotics, such as the aminoglycoside family of antibiotics, that bind RNA. In this review, the authors refer to the Mikkelsen reference also cited by the Examiner for this obviousness rejection. The Mikkelsen reference demonstrates the ability of aminoglycosides to interact with and inhibit the enzymatic activity of the RNA subunit of the *E. coli* RNase P holoenzyme. Antibiotics tend to bind promiscuously to RNAs and the Mikkelsen reference is directed to understanding how a particular family of antibiotics functions in relation to various specific RNA molecules. The Mikkelsen reference does not disclose a polypeptide having an RNase P consensus sequence (again, the *E. coli* RNase P is specifically excluded by claim 1), nor does it teach the use of an RNase P holoenzyme having such a polypeptide subunit to identify antibiotic agents.

The references cited by the Examiner describe the determination of the structure of a previously identified RNase P protein subunit and the modulation of several RNase P proteins by reagents such as aminoglycosides, puromycin, and cyclohexamide. Importantly, however, the RNase P holoenzymes described in these references do not include any of the RNase P polypeptides claimed in the present invention. The cited references also fail to describe a method using such polypeptides to identify an agent, which can be used as an antibiotic agent. The Potushak, Schroeder, Mikkelsen, and Spitzfaden references do not describe the polypeptides of the invention, nor do they describe the use of such polypeptides in a screen for antibiotic agents. Since the references do not describe the polypeptides of the invention, they cannot render the use of such polypeptides to identify an antibiotic agent obvious. These references do not teach or suggest all the claim limitations and therefore cannot render the claimed invention obvious. For this reason, the rejection should be withdrawn.

In addition, dependent claims 10 and 11 further specify that an antibiotic agent is identified using an assay for enzymatic activity of the RNase P holoenzyme that includes a fluorescently tagged $ptRNA^{Gln}$ or the use of fluorescence spectroscopy. Methods for labeling $ptRNA^{Gln}$ with a fluorescent tag and methods for measuring the enzymatic activity by fluorescence spectroscopy are described, for example, on pages 19 to 23 of the specification.

Nowhere is such a screening method disclosed or suggested by the cited references. The focus of Spitzfaden is the structure of the *S. aureus* RNase P protein and it does not include any sort of assay for RNase P activity. The RNase P assays disclosed by Potushak and Mikkelsen use a *radiolabeled RNase P substrate.* None of the references teach or suggest the use of a *fluorescently tagged oligonucleotide* for measuring the amount of either the remaining intact RNase P substrate or the cleaved substrate, as required by claims 10 and 11. As the cited

references do not render claim 10 or 11 obvious, this rejection of these claims should be withdrawn for these reasons as well as the reason cited above.

Rejection of claims 1, 2, 8-11, 13-21, and 29 under 35 U.S.C. § 103(a)

Claims 1, 2, 8-11, 13-21, and 29 are rejected, under 35 U.S.C. § 103(a), as being unpatentable over either of Spitzfaden (*J. Mol. Biol.* 295:105-115, 2000), Gopalan et al. (*J. Mol. Biol.*, 267:818-829, 1997) or Thompson et al. (*Nucleic Acids Research,* 22:4673-4680, 1994). The Examiner again states that the instant claims do not contain any rule as to what residues are considered essential, nor is there any indication of activity in SEQ ID NOs: 20-38. Applicants assert that, as stated in the arguments under the § 112, written description rejection, the rule for determining which polypeptides are RNase P protein subunits is defined in the specification, and in amended claim 1, as a sequence which conserves at least nine of the following twenty amino acids: R11, L12, F18, R46, G48, V51, K53, K54, A59, V60, R62, N63, K66, R67, R70, L80, D84, V86, L101, and L105. The results demonstrating the RNase P activity of polypeptides identified using the methods of the present invention are also discussed in detail above under the § 112 enablement rejections.

The standard for a prima facie case of obviousness has been set forth in *In re Vaeck, supra*, as described above. "To establish a prima facie case of obviousness....the prior art reference (or references when combined) must teach or suggest all the claim limitations." *In re Vaeck, supra.* This criteria is not satisfied in the present rejection of claims 1, 2, 8-11, 13-21, and 29 for obviousness.

Thompson describes parameters and modifications used to design an improved sequence alignment program. This reference describes general alignment methods and algorithms used for sequence alignments. Thompson does not in any way mention RNase P protein subunits or an alignment of RNase P protein subunits.

Spitzfaden describes the structure determination of the RNase P protein from *S. aureus* and an analysis of the RNA binding site on the protein surface. Spitzfaden includes an alignment of RNase P protein species from 11 known prokaryotic species. As stated in the attached Declaration of inventor Dr. Venkat Gopalan, identification of the nineteen sequences listed in Exhibit E as sequences of RNase P subunits was carried in the United States prior to January 7, 2000. The sequence alignment in Exhibit F of the nineteen RNase P protein sequences with sequences of previously known bacterial RNase P protein subunits was used in the identification of an RNase P consensus sequence. This consensus sequence includes R11, L12, F18, R46, G48, V51, K53, K54, A59, V60, R62, N63, K66, R67, R70, L80, D84, V86, L101, and L105 of *E. coli* RNase P. The identification of this RNase P consensus sequence was carried in the United States prior to the January 7, 2000 publication date of Spitzfaden. Because the claimed invention was reduced to practice prior to the publication of Spitzfaden, Spitzfaden cannot constitute prior art to the present claims under 35 U.S.C. § 102.

The above notwithstanding, Spitzfaden does not mention the use of a consensus sequence to identify novel RNAse P protein subunits, nor does Spitzfaden identify the 20 critical amino acids or the subset of 9 amino acids required by the present claims to identify an RNase P protein subunit.

Gopalan et al. describes the mutational analysis of the protein subunit of RNase P to examine the relevance of specific amino acids to the functional activity of the enzyme. Figure 2 of this reference shows the alignment of 12 known RNase P protein subunits used to identify common amino acids that may be relevant for enzymatic activity. This alignment does not identify the 20 amino acid consensus sequence that was identified in the present application, nor does it mention the relevance of the 9 amino acids that are also identified in the present application. In fact, as stated on page 825, the alignment revealed a "low degree of sequence

identity among the 12 RNase P protein sequences." The spatial conservation of the 9 out of 20 amino acids identified in the present application was also not suggested as the three-dimensional structure of RNase P had not yet been resolved (page 824). Not only does Gopalan et al. fail to identify or suggest the relevant 9 or 20 amino acids, Gopalan et al. also does not suggest the use of the common residues identified in the alignment to identify novel RNase P protein subunits.

The references, taken either individually or together, do not teach or suggest all of the limitations of the claims of the present invention. Therefore, the criteria for an obviousness rejection have not been met in the present case and this rejection should be withdrawn.

CONCLUSION

In summary, applicants submit that the claims are now in condition for allowance, and such action is respectfully requested.

Enclosed are a Petition to extend the period for replying to the final Office action for three months, to and including June 24, 2004, and a check in payment of the required extension fee.

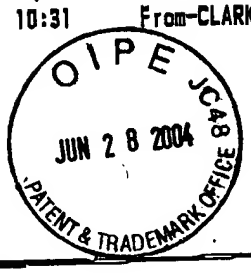
If there are any charges or any credits, please apply them to Deposit Account No. 03-2095.

Respectfully submitted,

Date: June 24, 2004

Kristina Bieker-Brady, Ph.D., P.C.
Reg. No. 39,109

Clark & Elbing LLP
101 Federal Street
Boston, MA 02110
Telephone: 617-428-0200
Facsimile: 617-428-7045



PATENT
ATTORNEY DOCKET NO. 50093/016001

Certificate of Mailing: Date of Deposit: June 24, 2004

I hereby certify under 37 C.F.R. § 1.8(a) that this correspondence is being deposited with the United States Postal Service as first class mail with sufficient postage on the date indicated above and is addressed to Mail Stop Amendment, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450

Susan M. Cannon
Printed name of person mailing correspondence

Susan M. Cannon
Signature of person mailing correspondence

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Venkat Gopalan et al.	Art Unit:	1652
Serial No.:	09/516,061	Examiner:	Charles L. Patterson Jr.
Filed:	March 1, 2000	Customer No.:	21559
Title:	NOVEL BACTERIAL RNASE P PROTEINS AND THEIR USE IN IDENTIFYING ANTIBACTERIAL COMPOUNDS		

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

DECLARATION OF DR. VENKAT GOPALAN UNDER 37 C.F.R. § 1.131

I, Venkat Gopalan, declare that:

1. I am an inventor of the invention described and claimed in the above-identified patent application.

2. In response to the Examiner's request for additional proof that there is only one RNase P polypeptide in each bacterial species, I submit data from the Clusters of Orthologous Groups of Proteins (COG) database (www.ncbi.nlm.nih.gov/cog). The COG database is a government supported database that provides a phylogenetic classification of the proteins encoded in complete genomes including bacterial, archael, and eukaryotic genomes. Each cluster group, or COG, is assigned a number and consists of an individual

protein from at least 3 lineages. The RNase P polypeptide subunit has been assigned COG594.

3. Using the Comparative Genomics web tool found at the web site for the Virtual Institute for Microbial Stress and Survival (VIMSS) supported by the U.S. Department of Energy Office of Science (http://vimss.lbl.gov/), I have searched the 131 genomes of the COG database for COG594.

4. This search produced 105 hits. Each of the bacteria listed in the 105 hits shows only one complete sequence for the RNase P polypeptide subunit. A copy of these results is attached as Exhibit D.

5. Included in the 105 hits are the RNase P polypeptide subunits for bacteria listed with multiple accession numbers in claim 1. Examples of these include *Streptomyces coelicolor* (A3(2)) and *Mycobacterium tuberculosis* (H37Rv).

6. I further certify that the amino acid sequences for the *N. gonorrhoaea* and *P. gingivalis* Rnase P polypeptide subunit used in the experiments described in my Declaration submitted on August 4, 2003 are identical to the sequences listed for the *N. gonorrhea* and *P. gingivalis* Rnase P polypeptide subunit in the specification as SEQ ID NOs: 27 and 31, respectively.

7. Claim 1, as amended in the present reply to Office action, recites polypeptides comprising an RNase P polypeptide consensus sequence having at least nine of the following twenty amino acids in the *E. coli* RNase P protein sequence: R11, L12, F18, R46, G48, V51, K53, K54, A59, V60, R62, N63, K66, R67, R70, L80, D84, V86, L101, and L105.

8. I, along with the other inventors, conceived of, and reduced to practice, the subject matter of claim 16 in the United States prior to January 7, 2000.

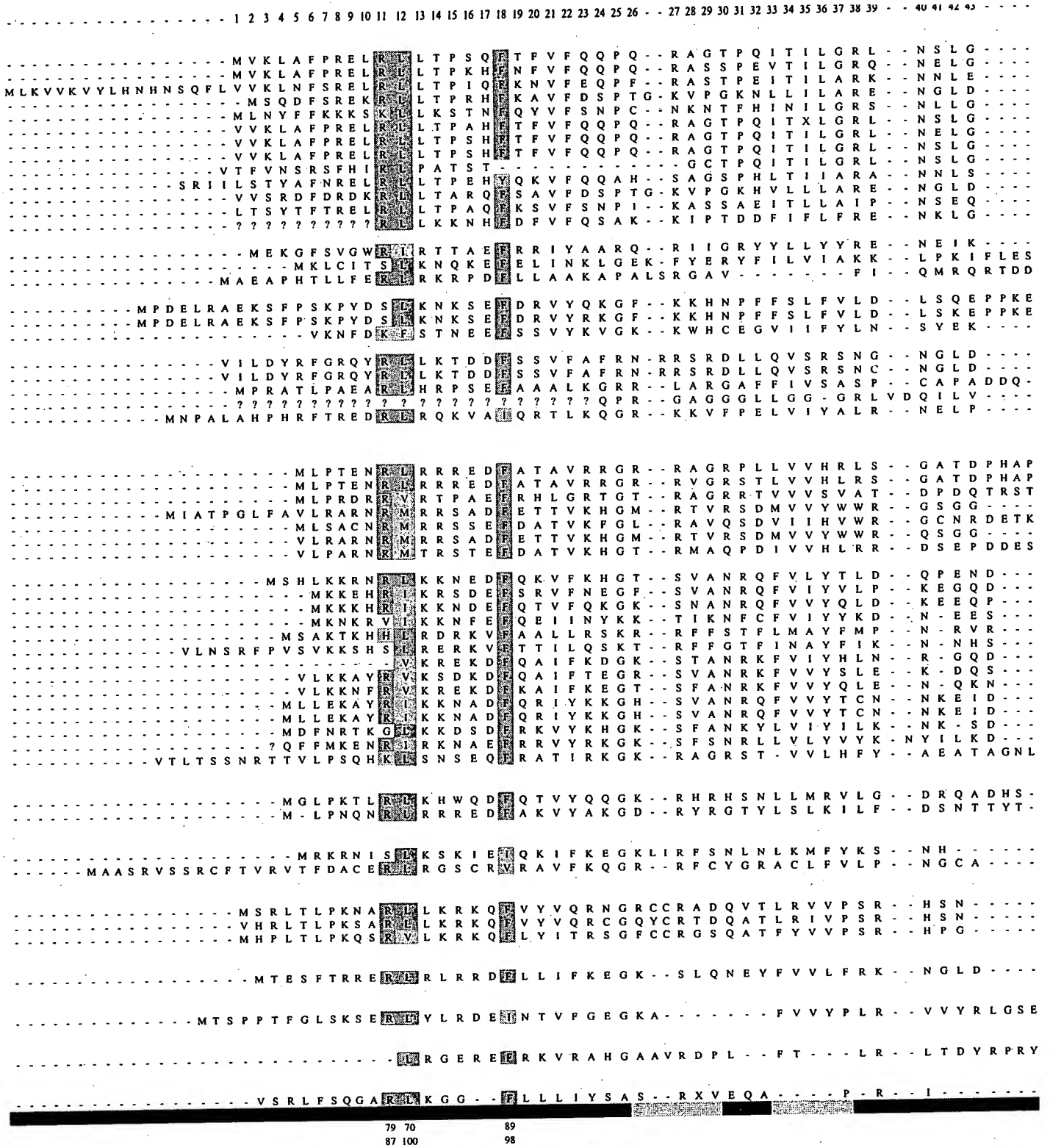
9. The reduction to practice of the claimed invention is evidenced by Exhibits E and F annexed hereto. The dates of electronic files containing Exhibit E and F are prior to January 7, 2000. Exhibit E contains the nucleic acid and translated amino acid sequences of nineteen RNase P subunits. The identification of these sequences as sequences of RNase P subunits was carried in the United States as described in the present application prior to January 7, 2000.

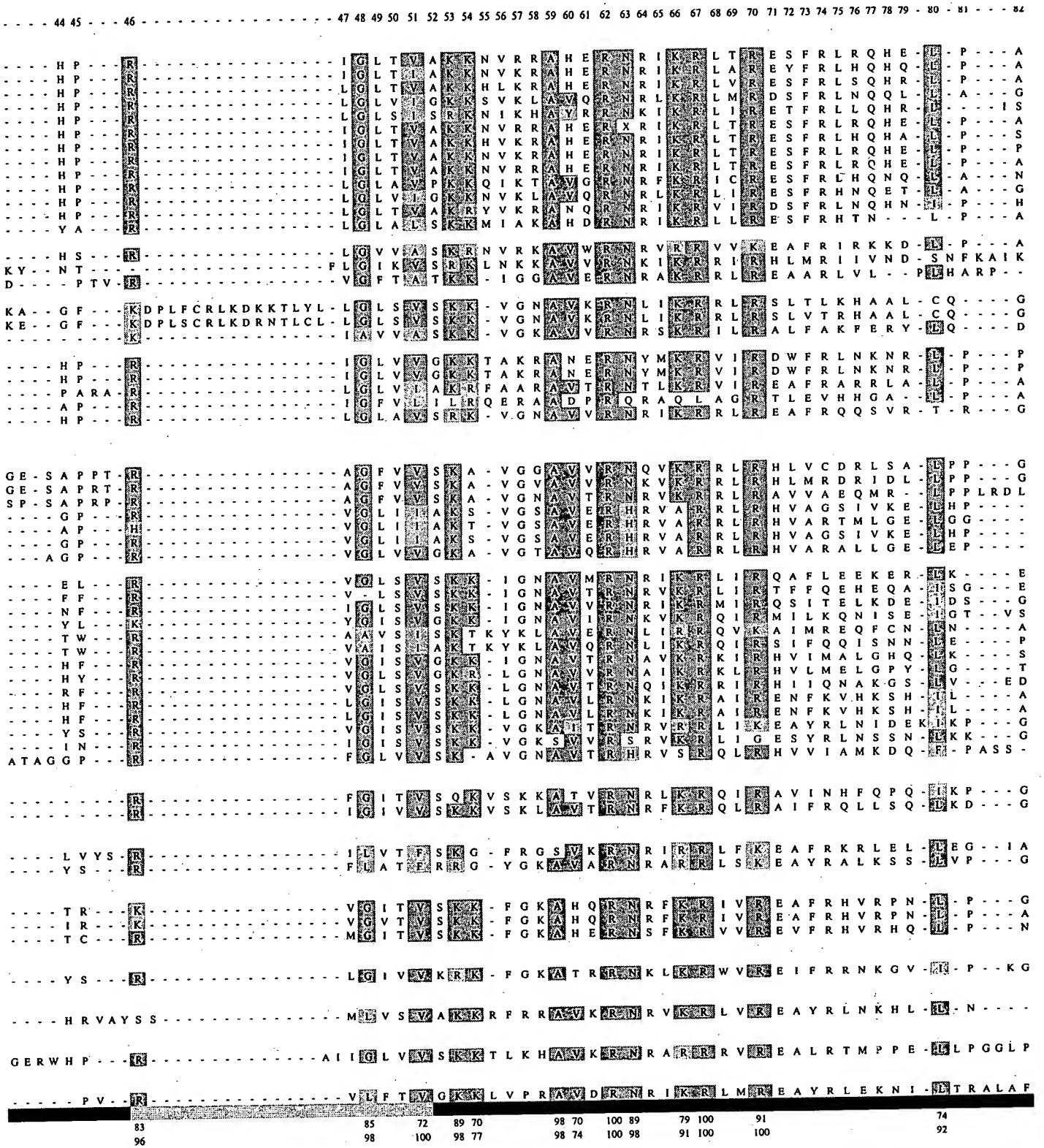
Exhibit F contains a sequence alignment of previously known bacterial RNase P protein subunits and RNase P sequences of the present invention using the ClustalW alignment program. Residues that were determined to be part of an RNase P consensus sequence are highlighted. These residues include R11, L12, F18, R46, G48, V51, K53, K54, A59, V60, R62, N63, K66, R67, R70, L80, D84, V86, L101, and L105 of *E. coli* RNase P. Thus, the identification of this RNase P consensus sequence was carried in the United States as described in the application prior to January 7, 2000.

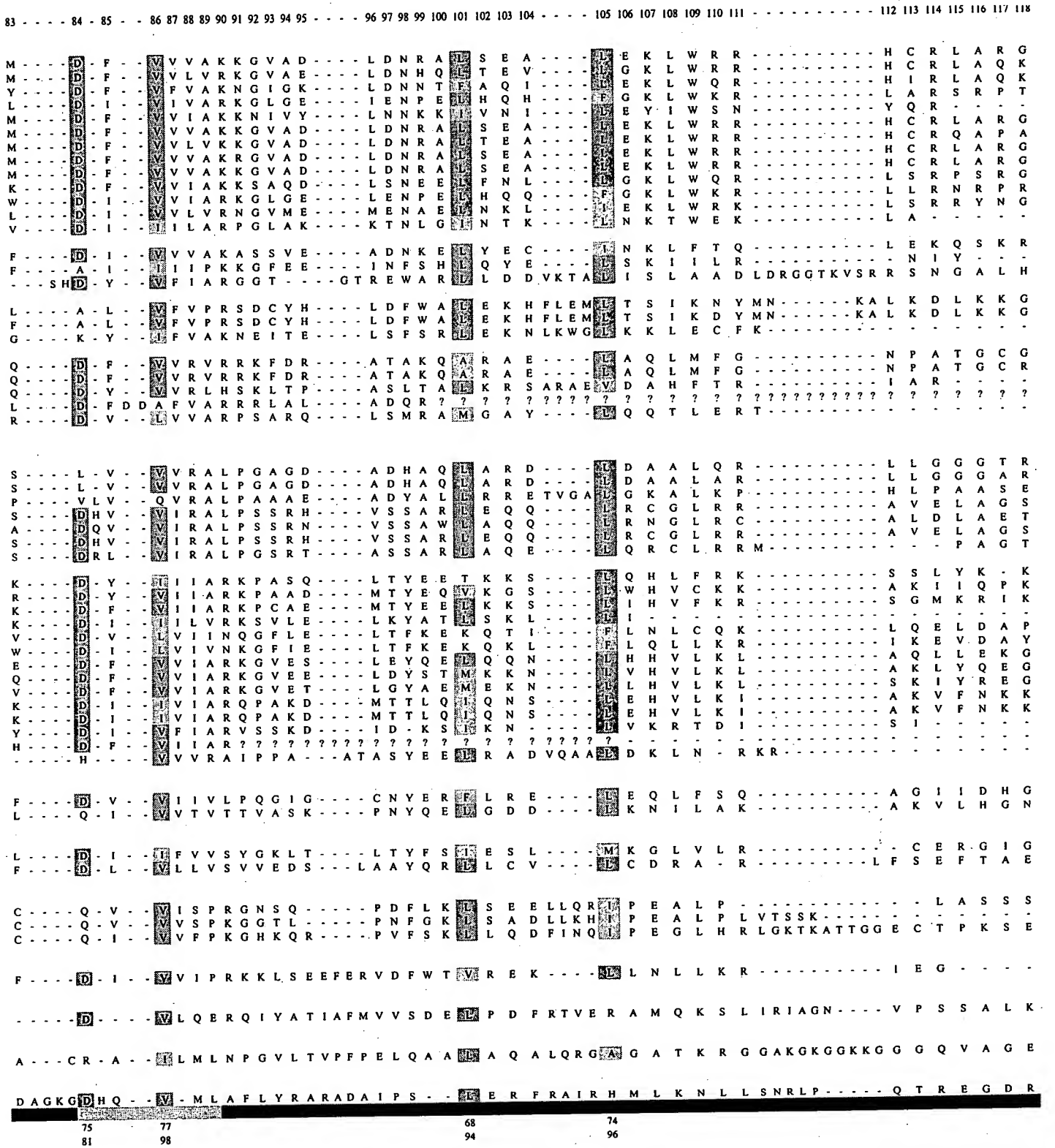
10. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patents issued thereon.

Date: 06-22-2004

Venkat Gopalan, Ph.D.
Assistant Professor
Department of Biochemistry
Ohio State University


Residue Number (Based on E. coli)
Gram Negative Bacteria
gamma purple
Escherichia coli (119)
Proteus mirabilis (119)
Haemophilus influenzae (136)
Pseudomonas putida (133)
Buchnera aphidicola (114)
Salmonella typhi (119)
Yersinia pestis (119)
Klebsiella pneumoniae* (119)
Salmonella paratyphi* (110)
Vibrio cholerae* (122)
Pseudomonas aeruginosa* (135)
Shewanella putrefaciens@ (118)
Legionella pneumophila& (101?)
alpha purple
Coxiella burnetii (121)
Rickettsia prowazekil (121)
Caulobacter crescentus@ (149)
epsilon purple
Helicobacter pylori 26695 (161)
Helicobacter pylori J99 (161)
Camphylobacter jejuni* (108)
beta purple
Neisseria gonorrhoeae* (123)
Neisseria meningitidis* (123)
Bordetella pertussis* (123)
Bordetella bronchiseptica& (787) seq. not complete
Thiobacillus ferrooxidans@ (116)
Gram Positive Bacteria
high G & C
Streptomyces bikiniensis (123)
Streptomyces coelicolor (123)
Micrococcus luteus (132)
Mycobacterium tuberculosis (125)
Mycobacterium leprae (120)
Mycobacterium bovis (115)
Mycobacterium avium* (119)
low G & C
Bacillus subtilis (119)
Bacillus halodurans (118)
Bacillus anthracis* (119)
Mycoplasma capricolum (102)
Mycoplasma pneumoniae (118)
Mycoplasma genitalium (128)
Streptococcus pyogenes* (113)
Streptococcus mutans* (119)
Streptococcus pneumoniae* (124)
Staphylococcus aureus NCTC* (117)
Staphylococcus aureus COL* (117)
Clostridium difficile* (114)
Clostridium acetobutylicum% (89?)
Corynebacterium diphtheriae* (129)
Cyanobacteria
Synechocystis PCC6803 (124)
Pseudanabaena PCC6903 (116)
Spirochaete
Borrella burgdorferi (119)
Treponema pallidum (133)
Chlamydiae
Chlamydia trachomatis (120)
Chlamydia trachomatis MoPn* (119)
Chlamydia pneumoniae (139)
Thermotoga
Thermotoga maritima (117)
Bacteroides
Porphyromonas gingivalis* (137)
Deinococci
Deinococcus radiodurans@ (166)
Green-Sulfur
Chlorobium tepidum@ (129)
Secondary Structure
% Identity
% Similarity
79 70
87 100
89
98


Residue Number (Based on E. coli)
Gram Negative Bacteria
gamma purple
Escherichia coli (119)
Proteus mirabilis (119)
Haemophilus influenzae (136)
Pseudomonas putida (133)
Buchnera aphidicola (114)
Salmonella typhi (119)
Yersinia pestis (119)
Klebsiella pneumoniae* (119)
Salmonella paratyphi* (110)
Vibrio cholerae* (122)
Pseudomonas aeruginosa* (135)
Shewanella putrefaciens@ (118)
Legionella pneumophila& (101?)
alpha purple
Coxiella burnetii (121)
Rickettsia prowazekii (121)
Caulobacter crescentus@ (149)
epsilon purple
Helicobacter pylori 26695 (161)
Helicobacter pylori J99 (161)
Camphylobacter jejuni* (108)
beta purple
Neisseria gonorrhoeae* (123)
Neisseria meningitidis* (123)
Bordetella pertussis* (123)
Bordetella bronchiseptica& (78?) seq. not complete
Thiobacillus ferrooxidans@ (116)
Gram Positive Bacteria
high G & C
Streptomyces bikiniensis (123)
Streptomyces coelicolor (123)
Micrococcus luteus (132)
Mycobacterium tuberculosis (125)
Mycobacterium leprae (120)
Mycobacterium bovis (115)
Mycobacterium avium* (119)
low G & C
Bacillus subtilis (119)
Bacillus halodurans (118)
Bacillus anthracis* (119)
Mycoplasma capricolum (102)
Mycoplasma pneumoniae (118)
Mycoplasma genitalium (128)
Streptococcus pyogenes* (113)
Streptococcus mutans* (119)
Streptococcus pneumoniae* (124)
Staphylococcus aureus NCTC* (117)
Staphylococcus aureus COL* (117)
Clostridium difficile* (114)
Clostridium acetobutylicum% (89?)
Corynebacterium diphtheriae* (129)
Cyanobacteria
Synechocystis PCC6803 (124)
Pseudanabaena PCC6903 (116)
Spirochaete
Borrelia burgdorferi (119)
Treponema pallidum (133)
Chlamydiae
Chlamydia trachomatis (120)
Chlamydia trachomatis MoPn* (119)
Chlamydia pneumoniae (139)
Thermotoga
Thermotoga maritima (117)
Bacteroides
Porphyromonas gingivalis* (137)
Deinococci
Deinococcus radiodurans@ (166)
Green-Sulfur
Chlorobium tepidum@ (129)
Secondary Structure
% Identity
% Similarity


Residue Number (Based on E. coli)
83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118
Gram Negative Bacteria
gamma purple
Escherichia coli (119)
Proteus mirabilis (119)
Haemophilus influenzae (136)
Pseudomonas putida (133)
Buchnera aphidicola (114)
Salmonella typhi (119)
Yersinia pestis (119)
Klebsiella pneumoniae* (119)
Salmonella paratyphi* (110)
Vibrio cholerae* (122)
Pseudomonas aeruginosa* (135)
Shewanella putrefaciens@ (118)
Legionella pneumophila& (101?)
alpha purple
Coxiella burnetii (121)
Rickettsia prowazekil (121)
Caulobacter crescentus@ (149)
epsilon purple
Helicobacter pylori 26695 (161)
Helicobacter pylori J99 (161)
Camphylobacter jejuni* (108)
beta purple
Neisseria gonorrhoeae* (123)
Neisseria meningitidis* (123)
Bordetella pertussis* (123)
Bordetella bronchiseptica& (787) seq. not complete
Thiobacillus ferrooxidans@ (116)
Gram Positive Bacteria
high G & C
Streptomyces bikiniensis (123)
Streptomyces coelicolor (123)
Micrococcus luteus (132)
Mycobacterium tuberculosis (125)
Mycobacterium leprae (120)
Mycobacterium bovis (115)
Mycobacterium avium* (119)
low G & C
Bacillus subtilis (119)
Bacillus halodurans (118)
Bacillus anthracis* (119)
Mycoplasma capricolum (102)
Mycoplasma pneumoniae (118)
Mycoplasma genitalium (128)
Streptococcus pyogenes* (113)
Streptococcus mutans* (119)
Streptococcus pneumoniae* (124)
Staphylococcus aureus NCTC* (117)
Staphylococcus aureus COL* (117)
Clostridium difficile* (114)
Clostridium acetobutylicum% (89?)
Corynebacterium diphtheriae* (129)
Cyanobacteria
Synechocystis PCC6803 (124)
Pseudanabaena PCC6903 (116)
Spirochaete
Borrelia burgdorferi (119)
Treponema pallidum (133)
Chlamydiae
Chlamydia trachomatis (120)
Chlamydia trachomatis MoPn* (119)
Chlamydia pneumoniae (139)
Thermotoga
Thermotoga maritima (117)
Bacteroides
Porphyromonas gingivalis* (137)
Deinococci
Deinococcus radiodurans@ (166)
Green-Sulfur
Chlorobium tepidum@ (129)
Secondary Structure
% Identity
% Similarity
75
81
77
98
68
94
74
96

Residue Number (Based on *E. coli*)	119
Gram Negative Bacteria	
gamma purple	
Escherichia coli (119)	S
Proteus mirabilis (119)	S
Haemophilus influenzae (136)	S
Pseudomonas putida (133)	P A V T A N S A G V D S Q D A
Buchnera aphidicola (114)	-
Salmonella typhi (119)	S
Yersinia pestis (119)	S
*Klebsiella pneumoniae** (119)	S
*Salmonella paratyphi** (110)	S
*Vibrio cholerae** (122)	-
*Pseudomonas aeruginosa** (135)	T E S P A D A P G V A D G T H A
Shewanella putrefaciens@ (118)	-
Legionella pneumophila& (101?)	-
alpha purple	
Coxiella burnetii (121)	S S S V
Rickettsia prowazekii (121)	-
Caulobacter crescentus@ (149)	D A A P S S Q P D P T V S G
epsilon purple	
Helicobacter pylori 26695 (161)	I T H T Y A K Q
Helicobacter pylori J99 (161)	M T H T H A K Q
*Camphylobacter jejuni** (108)	-
beta purple	
Neisseria gonorrhoeae* (123)	K Q V
*Neisseria meningitidis** (123)	K Q A
*Bordetella pertussis** (123)	-
Bordetella bronchiseptica& (78?) seq. not complete	? ? ?
Thiobacillus ferrooxidans@ (116)	-
Gram Positive Bacteria	
high G & C	
Streptomyces bikiniensis (123)	-
Streptomyces coelicolor (123)	-
Micrococcus luteus (132)	H A
Mycobacterium tuberculosis (125)	D R
Mycobacterium leprae (120)	D W
Mycobacterium bovis (115)	D
*Mycobacterium avium** (119)	G P
low G & C	
Bacillus subtilis (119)	S S S K
Bacillus halodurans (118)	V R A H K
*Bacillus anthracis** (119)	S S V R K
Mycoplasma capricolum (102)	-
Mycoplasma pneumoniae (118)	K P K
Mycoplasma genitalium (128)	Q T S A N K
*Streptococcus pyogenes** (113)	F E S E E K H
*Streptococcus mutans** (119)	S I R E K E
*Streptococcus pneumoniae** (124)	N G S E K E T K V D
Staphylococcus aureus NCTC* (117)	I K
Staphylococcus aureus COL* (117)	I K
*Clostridium difficile** (114)	-
Clostridium acetobutylicum% (89?)	-
*Corynebacterium diphtheriae** (129)	-
Cyanobacteria	
Synechocystis PCC6803 (124)	H S R N H L L
Pseudanabaena PCC6903 (116)	-
Spirochaete	
Borrelia burgdorferi (119)	E S K
Treponema pallidum (133)	R Q E
Chlamydiae	
Chlamydia trachomatis (120)	R C
Chlamydia trachomatis MoPn* (119)	-
Chlamydia pneumoniae (139)	K C V T A P R
Thermotoga	
Thermotoga maritima (117)	-
Bacteroides	
*Porphyromonas gingivalis** (137)	N E
Deinococci	
Deinococcus radiodurans@ (166)	R A G N E S G S G R V S A E V N P T S A S P A S P A E K S
Green-Sulfur	
Chlorobium tepidum@ (129)	I E
Secondary Structure	
% Identity	
% Similarity	

J. Mol. Biol. (1997) 267, 818–829

JMB

Analysis of the Functional Role of Conserved Residues in the Protein Subunit of Ribonuclease P from *Escherichia coli*

Venkat Gopalan[1], Andreas D. Baxevanis[2], David Landsman[2] and Sidney Altman[1]*

[1]*Department of Biology, Yale University, New Haven, CT 06520-8103, USA*

[2]*National Center for Biotechnology Information National Library of Medicine National Institutes of Health Room 8N-805, Building 38A Bethesda, MD 20894, USA*

**Corresponding author*

The processing of precursor tRNAs and some other small cellular RNAs by M1 RNA, the catalytic subunit of *Escherichia coli* ribonuclease P, is accelerated by C5 protein (the protein cofactor) both *in vitro* and *in vivo*. In an effort to understand the mechanism by which the protein cofactor promotes and stabilizes certain conformations of M1 RNA that are most efficient for RNase P catalysis, we have used site-directed mutagenesis to generate mutant derivatives of C5 protein and assessed their ability to promote RNase P catalysis *in vivo* and *in vitro*. Our results indicate that certain conserved hydrophobic and basic residues in C5 protein are important for its function and that single amino acid residue changes in C5 protein can alter the substrate specificity of the RNase P holoenzyme.

Keywords: RNase P protein subunit; site-directed mutagenesis

Introduction

Ribonuclease P (RNase P) is an endoribonuclease that cleaves the 5′-terminal leader sequences of precursor tRNAs (ptRNAs: Figure 1; Altman *et al.*, 1995; Pace & Brown, 1995). In addition to ptRNAs, RNase P from *Escherichia coli* can cleave other endogenous substrates such as the precursors to 4.5S RNA and 10Sa RNA (Bothwell *et al.*, 1976; Komine *et al.*, 1994). RNase P (in conjunction with RNase E) is also involved in the processing of the polycistronic mRNA of the histidine operon in *Salmonella typhimurium* (Alifano *et al.*, 1994). The RNase P holoenzyme of *E. coli* consists of a catalytic RNA subunit (M1 RNA, 377 nucleotides) and a protein subunit (C5, 119 amino acid residues). Under certain conditions *in vitro*, M1 RNA can catalyze the hydrolysis of ptRNAs even in the absence of C5 protein; however, both M1 RNA and C5 protein are essential for the activity of RNase P *in vivo*. In contrast to the RNA subunits of RNase P from Bacteria, those of RNase P from Archaea and Eukarya fail to exhibit catalytic activity *in vitro* in the absence of their protein subunits.

Studies performed with different substrates and M1 RNA, in either the presence or absence of C5 protein, have revealed that M1 RNA is a more efficient and versatile enzyme in the presence of the protein cofactor (Guerrier-Takada *et al.*, 1983; Lumelsky & Altman, 1988; Reich *et al.*, 1988; Peck-Miller & Altman, 1991). Determination of kinetic parameters for the hydrolysis of numerous substrates by M1 RNA alone and the RNase P holoenzyme has revealed that M1 RNA (by itself) can achieve the most efficient conformation for recognition of some substrates, while for some others it requires the C5 protein (Kirsebom & Altman, 1989; Peck-Miller & Altman, 1991; Kirsebom & Svärd, 1992). Furthermore, the presence of C5 protein can alleviate the deleterious effect of mutations on the activity of M1 RNA in different parts of the M1 RNA molecule (Lumelsky & Altman, 1988). Therefore, C5 protein must engage in specific interations with the catalytic RNA subunit to stabilize certain conformations of M1 RNA favorable for catalysis and thus play a critical role in recognition/binding of some substrates by the RNase P holoenzyme.

The amino acid residues in C5 protein responsible for the various effects on M1 RNA catalysis have not been identified. Here we have used site-directed mutagenesis to examine the role of conserved residues in C5 protein with regard to RNase P activity *in vivo* and *in vitro*.

Present address: A. D. Baxevanis, Computational Genomics Core, Genome Technology Branch, National Center for Genome Research, National Institutes of Health, Bethesda, MD 20892, USA.

Abbreviations used: RNase P, ribonuclease P; ptRNAs, precursor tRNAs; ts, temperature sensitive; IPTG, isopropyl-β-D-thiogalactopyranoside; p4.5S, precursor to 4.5S RNA; ss, single-stranded.

0022–2836/97/140818–12 $25.00/0/mb970906

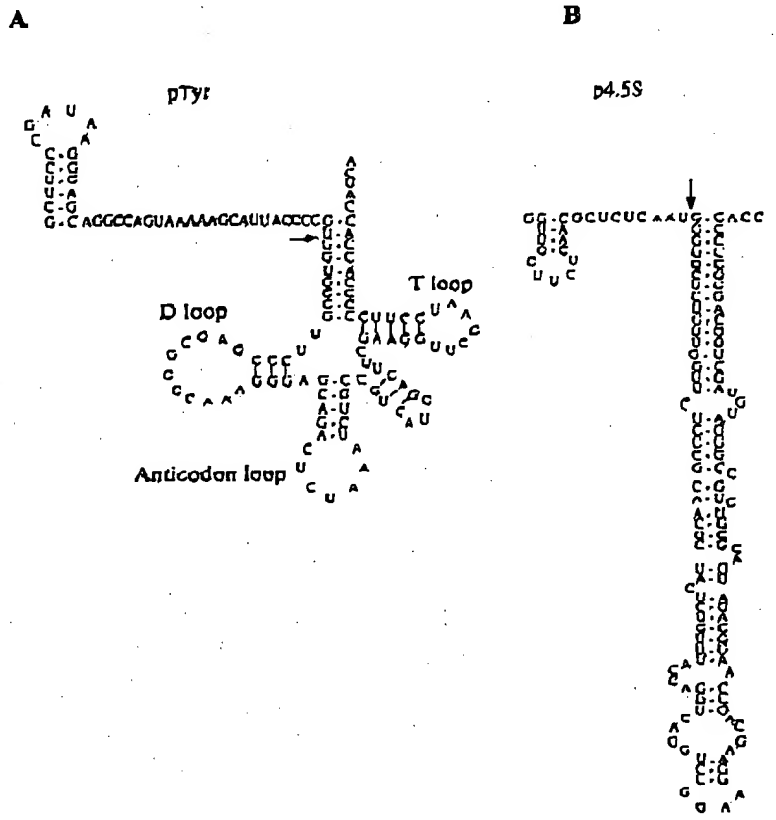


Figure 1. Secondary structure of $ptRNA^{Tyr}$ and p4.5S RNA, substrates of *E. coli* RNase P. The site of RNase P-mediated cleavage in these substrates is indicated by arrowheads.

Results

Rationale

There are a few conserved basic and hydrophobic residues which contribute to the limited identity observed among different prokaryotic RNase P protein subunits (Figure 2). However, there are several positions in the polypeptide chain at which the physicochemical properties of the residues are similar. None of the consensus RNA-binding motifs (Burd & Dreyfuss, 1994) identified in other RNA-binding proteins is present in C5 protein. However, since arginine-rich sequences in bacteriophage, viral, and ribosomal proteins are thought to mediate RNA recognition (Lazinski *et al.*, 1989; Burd & Dreyfus, 1994), it is noteworthy that there is a stretch of ten residues in C5 protein (from position 60 to 70) that is rich in arginine and lysine residues.

The affinity of C5 protein for M1 RNA increases 500-fold as the ionic strength is increased from 0.1 to 1.0 M NH_4Cl (Talbot & Altman, 1994). The salt dependence of the C5 protein-M1 RNA interaction suggests that hydrophobic interactions play a role in holoenzyme formation. Also, recent crystallographic studies provide evidence for aromatic amino acid residues in RNA-binding proteins stacking on RNA bases in their RNA ligands (Oubridge *et al.*, 1994; Nagai, 1996). Therefore, in this study, emphasis has been placed on altering aromatic residues in addition to certain basic residues in C5 protein.

Genetic complementation

The mutation R46H in the chromosomal gene encoding C5 protein results in a temperature-sensitive (ts) phenotype in *E. coli* (Schedl & Primakoff, 1973; Kirsebom *et al.*, 1988). Recently, an *E. coli* strain, *T7A49*, which contains both the *rnpA*49 mutation (i.e. C5 R46H) and the T7 RNA polymerase gene in its chromosome was constructed (Guerrier-Takada *et al.*, 1995). Transformation of these cells with a plasmid bearing the gene encoding wild-type C5 protein can rescue this mutation and abolish the ts phenotype. The various mutant derivatives of C5 protein (subcloned in the same vector as the wild-type C5 protein and under control of the T7 RNA polymerase promoter) were analyzed for their ability to complement the R46H mutation in *T7A49* cells. In λ DE3 lysogens (such as *T7A49*

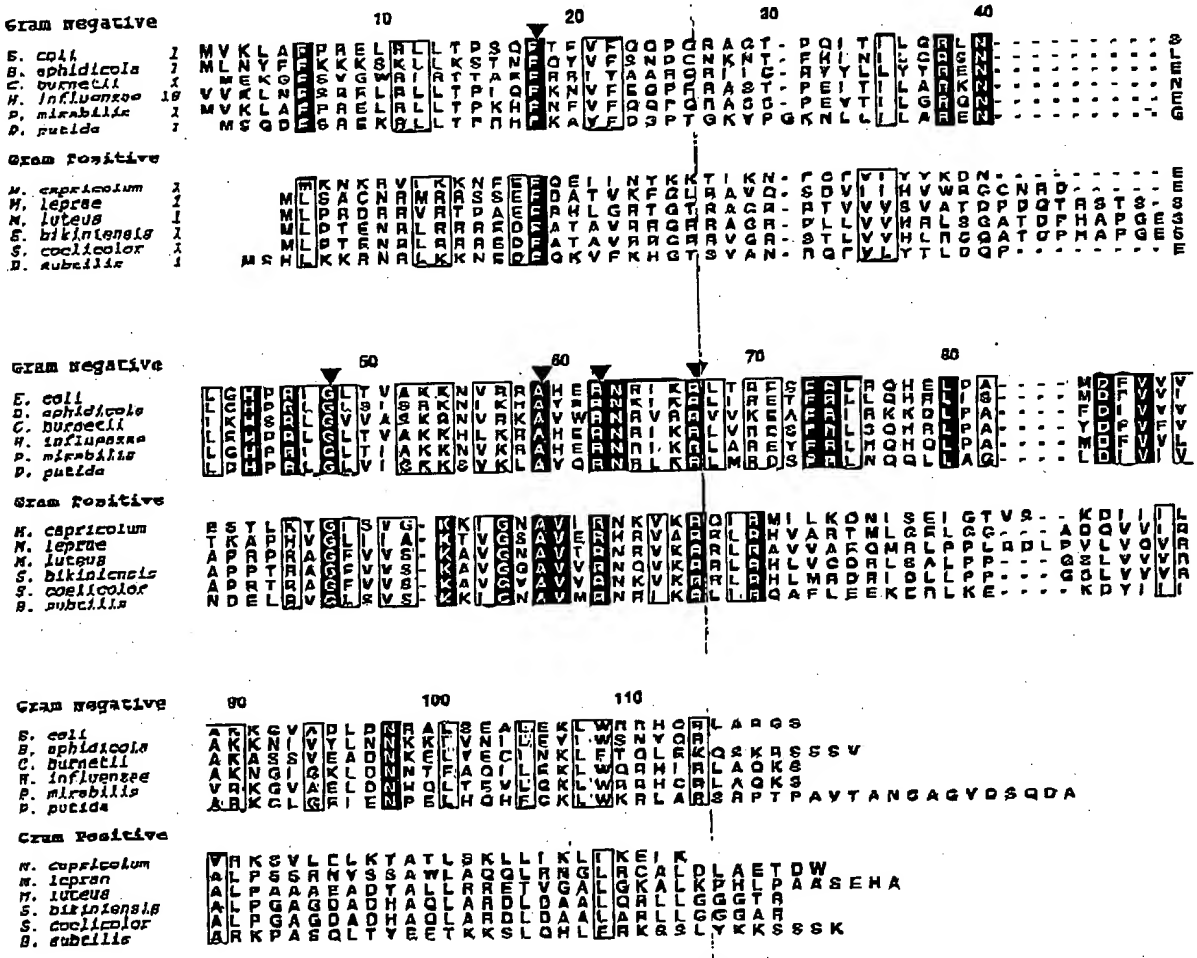


Figure 2. Multiple alignment of sequences of protein subunits of RNase P from Bacteria using Clustal W (Thompson *et al.*, 1994). Sequences are grouped as Gram negative and Gram positive. The numbering scheme at the top of the Figure is based on the sequence of the protein subunit of *E. coli* RNase P, and the numbers preceding each sequence indicate the position of the first shown residue for that sequence. Identical and conserved regions were calculated for each group. Vertical arrowheads identify positions in the alignment where the residues for all the sequences are identical. Positions with absolute conservation have a dark background, while those positions showing only conservative substitutions have a gray background. The sequences of the protein subunits of RNase P from *Escherichia coli, Buchnera aphidicola, Coxiella burnetii, Haemophilus influenzae, Proteus mirabilis, Pseudomonas putida, Mycoplasma capricolum, Mycobacterium leprae, Micrococcus luteus, Streptomyces bikiniensis, Streptomyces coelicolor,* and *Bacillus subtilis* were used in this alignment.

cells) the T7 RNA polymerase gene is under the control of an isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible *lac* UV5 promoter, which allows a low level of transcription even in the uninduced state. Therefore, our complementation analyses were performed in the absence of IPTG.

The results of our complementation analyses are summarized in Table 1. In general, altering the identity of the conserved residues exerts a moderate to severe effect on the activity of the protein *in vivo*. While substitution of some of the conserved hydrophobic residues individually with Ala had only a moderate effect on their function *in vivo*, changing two hydrophobic residues simultaneously to Ala resulted in a very drastic effect. For example, the single mutants C5 F22A, C5 F73A, and C5 W109A can partially rescue the ts phenotype of the *T7A49* cells while the double mutants C5 F22A/W109A and C5 F73A/W109A fail to exhibit any activity *in vivo* (Table 1). The alteration of positively charged residues (such as R62A or R67A) had a moderate effect on the ability of C5 protein to rescue the ts phenotype of *T7A49* cells. Here we have not constructed any double mutants in which two basic residues were altered simultaneously. Also, the mutants C5 R57P and C5

Table 1. Results of RNase P assays using C5 protein or its mutant derivatives

Protein used to reconstitute with M1 RNA	ptRNA^{Tyr} Relative initial velocity (%) 30°C	ptRNA^{Tyr} Relative initial velocity (%) 43°C	p4.5S RNA Relative activity (%) 43°C	*T7A49* Complementation* 43°C
Wild-type C5	100	100	100	++
F18A	36	34	2	ND
F22A	42	41	<1	+
F73A	76	72	64	+
W109A	76	69	44	+
F18W, W109F	64	87	48	+
F18A, W109A	<1	<1	<1	−
F18A, F22A	7	3	<1	ND
F22A, W109A	43	8	<1	−
F18A, F73A	13	2	<1	−
F73A, W109A	2	<1	<1	−
R62A	81	67	<1	+
N63V	71	84	36	++
K66A	55	50	14	+
R67A	49	18	4	+

* Mutants have been classified qualitatively into three categories based on their ability to rescue the ts phenotype of *T7A49* cells grown in liquid media. If the complementation observed with the mutant protein was comparable to that of the wild-type protein, the mutants were classified as ++; the mutants which behaved like the untransformed *T7A49* cells and showed a complete loss of growth after two hours of heat shock were classified as −. There is an intermediate classification defined as +, which refers to mutants that were able to either weakly or moderately rescue the ts phenotype. This classification into three groups, although arbitrary, has enabled us to determine if mutations introduced in C5 protein had a severe, moderate, or nil effect on the activity of the protein *in vivo*. ND, not determined.

N63P were not able to rescue the ts phenotype of *T7A49* cells (data not shown).

RNase P assays using C5 protein or its mutant derivatives

We examined the ability of the various mutants to participate in RNase P catalysis *in vitro*. Wild-type C5 and its mutant derivatives were purified subsequent to their overexpression in an *E. coli* strain, *BL21 (DE3)*, a λ lysogen in which expression of the T7 RNA polymerase gene is regulated by the *lac* UV5 promoter. Fractionation of the crude cell extracts revealed that many of the overexpressed mutant proteins were present (to varying extents) in the P30 (the pellet obtained after centrifugation at 30,000*g*) unlike the wild-type C5 protein, which fractionates to the S30 (the supernatant obtained after centrifugation at 30,000*g*). It is possible that these mutant derivatives of C5 protein, when overexpressed in *BL21 (DE3)* cells at 37°C, aggregate and form inclusion bodies. A purification procedure has been described by Baer *et al.* (1989) for isolating C5 R46H from the P30. We have employed the same procedure (with some modifications) to purify the various mutant derivatives that were constructed in this study (see Materials and Methods). Although the purification procedure involves the use of a strong denaturing agent, we have verified using spectroscopic techniques that several of these mutants thus isolated do regain structure after the step-wise removal of urea (Gopalan *et al.*, 1997). All the mutant derivatives of C5 protein were purified to near homogeneity as judged by silver staining of SDS-polyacrylamide gels.

The various mutant derivatives of C5 protein were reconstituted with wild-type M1 RNA and the ability of these holoenzymes to cleave the precursors to tRNA^{Tyr} (ptRNA^{Tyr}) and 4.5S RNA (p4.5S RNA) was examined. A vast excess (200-fold) of the protein relative to M1 RNA was used to ensure that holoenzyme (M1 RNA + C5 protein) assembly was favored even with those mutant derivatives of C5 protein that might exhibit RNA-binding defects. We performed RNase P assays at 30°C and 43°C to check for thermosensitivity. The results of these assays are depicted in Figure 3 (for cleavage of ptRNA^{Tyr}) and Figure 4 (for cleavage of p4.5S RNA).

Altering amino acid residues in C5 protein elicits various effects on RNase P catalysis. There are mutants (such as C5 F18A, C5 F22A, and C5 R62A) which can cleave ptRNA^{Tyr} fairly efficiently but not p4.5S RNA (compare lanes 3, 4 and 14 in Figure 3A *versus* Figure 4A). Some mutants display a decrease in activity with increase in assay temperature from 30°C to 43°C (for example, C5 F22A/W109A; compare lane 10 in Figure 3A *versus* 3B). There are at least two mutants (C5 F18A/W109A and C5 F73A/W109A) which are severely defective in promoting RNase P activity with either substrate at both 30°C and 43°C (Figure 3A and B, lanes 8 and 12).

The initial velocity for hydrolysis of ptRNA^{Tyr} was measured for the mutant holoenzymes and compared with that of the wild-type holoenzyme (Table 1). The initial velocity observed for the wild-

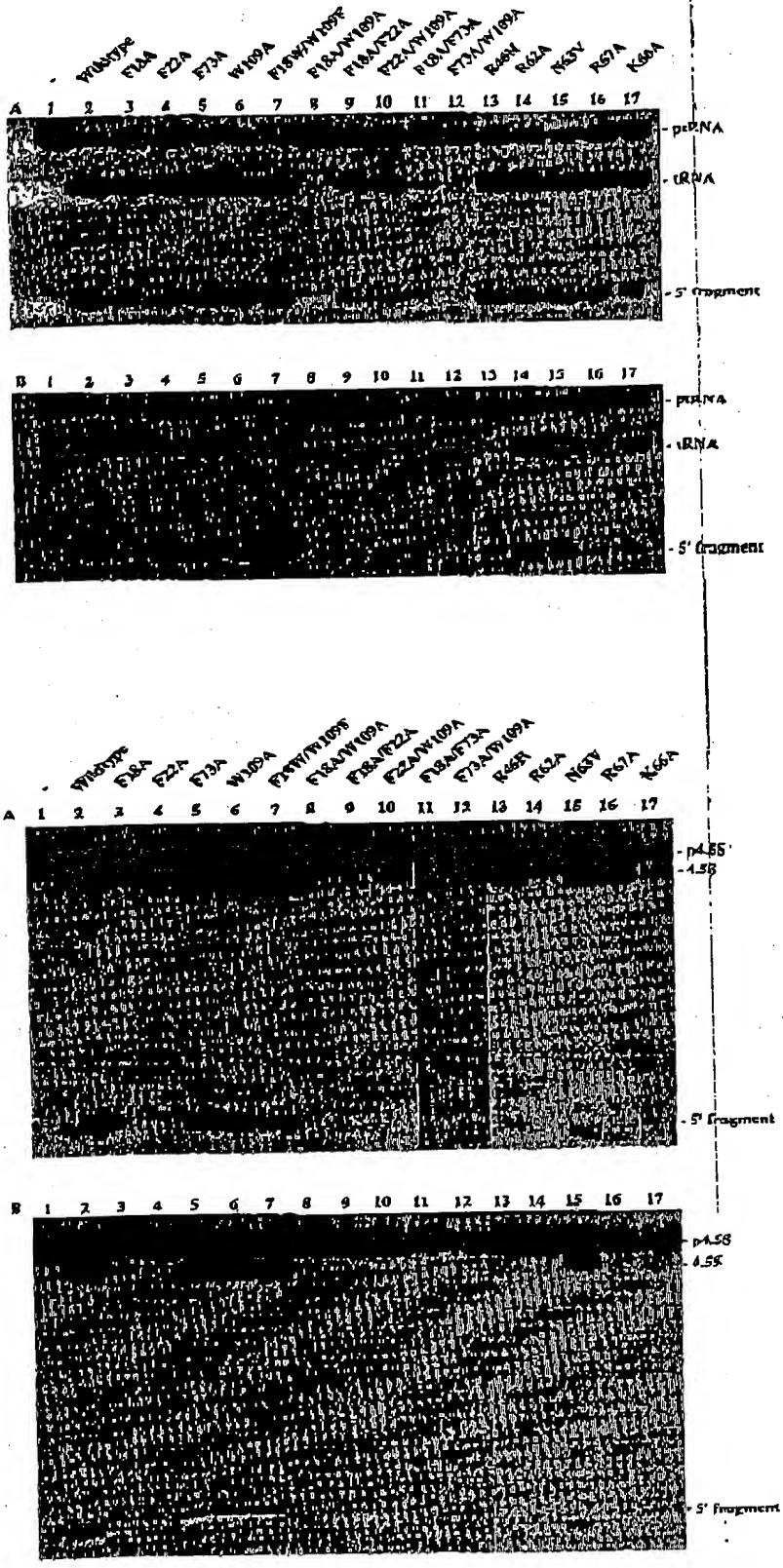


Figure 3. Effect of mutations in C5 protein on the activity of the RNase P holoenzyme with ptRNA^{Tyr} as the substrate. Holoenzymes, composed of M1 RNA (1 nM) and either wild-type C5 protein or its mutant derivatives (200 nM), were reconstituted and then assayed for activity at either 30°C (A) or 43°C (B). For more details, refer to Materials and Methods.

Figure 4. Effect of mutations in C5 protein on the activity of the RNase P holoenzyme with p4.5S RNA as the substrate. Holoenzymes, composed of M1 RNA (1 nM) and either wild-type C5 protein or its mutant derivatives (200 nM), were reconstituted and then assayed for activity at either 30°C (A) or 43°C (B). For more details, refer to Materials and Methods.

type holoenzyme at 30°C and 43°C are 12 min^{-1} and 36 min^{-1}, respectively. The initial velocities observed with the mutant holoenzymes are reported as the relative activities compared to that of the wild-type holoenzyme (Table 1). While alteration of some of the conserved hydrophobic residues (for example, F18A, F22A, F73A, and W109A) individually resulted in a maximum of threefold decrease in RNase P activity, changing two hydrophobic residues simultaneously to Ala rendered the mutants nearly inactive, especially when the assay temperature was 43°C. This pattern is reminiscent of the results obtained with the complementation assay described above.

For the hydrolysis of p4.5S RNA by the wild-type and mutant RNase P holoenzymes, the initial velocity was not calculated. The relative activity of the mutant holoenzymes, compared to the wild-type holoenzyme, is provided in Table 1 and was calculated on the basis of per cent cleavage observed while assaying for activity with 100 nM substrate at 43°C for ten minutes. There were no qualitative differences in the relative activities when the same assay was performed with 500 nM substrate (data not shown).

Discussion

Effect of mutations in C5 protein on RNase P activity *in vitro*

We have demonstrated that certain hydrophobic and basic residues in C5 protein are important for RNase P catalysis *in vivo* and *in vitro* and that some of the single amino acid changes in C5 protein alter the substrate specificity of the RNase P holoenzyme.

The impairment of function (i.e. RNase P catalysis) observed with certain mutant derivatives of C5 protein could result from: (1) the inability of the mutants to fold into a stable tertiary structure, or (2) loss of critical nucleic acid/amino acid residue contacts in the holoenzyme complex. Alteration of side-chains that are fully or partially buried in the wild-type C5 protein structure might result in either local or global destabilization of the tertiary structure depending on: (1) the ability of the protein to adapt to this mutation (*via* structural rearrangements), and (2) the extent to which the cavity, created as a result of the mutation, is deleterious to function (Matthews, 1993; Cordes *et al.*, 1996). The substitution of solvent-exposed amino acid residues (for example, Arg62 or Lys66) with Ala would *a priori* not be expected to perturb the tertiary structure of the protein. Since solvent-exposed residues form the molecular surfaces that mediate binding to ligands, the alteration of such residues in C5 protein might result in loss of contacts in the holoenzyme complex.

The observed increase in affinity of C5 protein for M1 RNA with increasing ionic strength suggests that hydrophobic interactions play a role in holoenzyme assembly (Talbot & Altman, 1994). The effects observed with the various mutant derivatives of C5 protein in which hydrophobic residues have been altered individually, or in pairs, implicate these residues as being important for holoenzyme assembly and function. The replacement of conserved hydrophobic residues individually was much less detrimental to function than the simultaneous alteration of two hydrophobic residues. For instance, when $ptRNA^{Tyr}$ served as the substrate, the holoenzymes reconstituted with mutants C5 F18A, C5 W109A, and C5 F18A/W109A exhibit initial velocities of 34%, 69% and <1%, respectively, relative to that of the wild-type holoenzyme (Table 1). On the basis of fluorescence spectroscopic analyses performed with wild-type C5 protein and its mutant derivatives C5 F18A, C5 F73A, C5 F22A, C5 F18A/F73A, C5 F18A/F22A, and C5 F18W/W109F, we concluded that Phe18 and Phe73 influence the fluorescence emission of Trp109 and that Phe18 and Phe73 are proximal to Trp109 in the tertiary structure as part of an aromatic core (Gopalan *et al.*, 1997). It is noteworthy that the two mutants C5 F18A/W109A and C5 F73A/W109A display negligible RNase P activity at 30°C and 43°C with the two substrates used in this study (Figures 3 and 4) and certainly are the most severely compromised mutants with regard to RNase P catalysis. The fluorescence data taken together with our functional analysis of the various mutants suggest that Phe18, Phe73, Phe22, and W109 play an important role in maintaining the structural core of the protein. Studies are in progress to determine the stability of the various mutant derivatives using circular dichroism spectroscopy.

There are some hydrophobic mutants which display a decrease in activity with increasing temperature. For example, the relative activities (compared to that of the wild-type holoenzyme) observed with the mutants C5 F18A/F73A and C5 F22A/W109A are 13% and 43%, respectively, at 30°C, and 2% and 8%, respectively, at 43°C (Table 1). This result suggests that destabilization caused by certain hydrophobic mutations manifests in a defective phenotype only at the higher temperature, perhaps a reflection of the adaptability of the protein to these alterations at the lower temperature. This is consistent with results from various studies on the structural responses of proteins to replacement of amino acid residues, which have revealed an unexpected degree of tolerance to even seemingly disruptive hydrophobic mutations (Matthews, 1993; Cordes *et al.*, 1996).

The mutants C5 R62A, C5 K66A, and C5 R67A have enabled us to examine the role of some conserved basic residues in RNase P catalysis. Although mutants C5 R62A and C5 K66A are able to efficiently cleave $ptRNA^{Tyr}$, their ability to cleave p4.5S RNA is compromised. Unlike C5 R62A and C5 K66A, which do not display a thermosensitive phenotype with regard to $ptRNA^{Tyr}$ cleavage, C5 R67A displays higher activity at 30°C (with both substrates tested here) compared to that

observed at 43°C. This behavior of C5 R67A is reminiscent of that of C5 R46H reported by Baer *et al.* (1989). The catalytic efficiency of the C5 R46H mutant holoenzyme was lower than that of the wild-type holoenzyme and it was demonstrated that assembly of the mutant holoenzyme was defective (Baer *et al.*, 1989). Multiple weak, non-covalent bonds play a crucial role in stabilizing macromolecular complexes such as the RNase P holoenzyme. The absence of such interactions in the mutant RNase P holoenzyme complexes, containing either C5 R67A or C5 R46H, will presumably manifest as a thermosensitive phenotype of functional activity on account of the low stability of the holoenzyme.

The alteration of conserved residues in mutants such as C5 N63V and C5 F73A does not result in drastic changes in the *in vitro* activity of the protein (under conditions examined here). Furthermore, for the purpose of RNA-protein footprinting experiments, we have prepared some single cysteine-substituted mutants (for example, C5 S16C/C113S) that are active both *in vivo* and *in vitro*. All these mutants serve as controls to indicate that not all mutations are detrimental to the function of the protein (data not shown).

Altered substrate specificity of certain mutant RNase P holoenzymes

The mutants C5 F18A, C5 F22A, and C5 R62A are fairly efficient in promoting hydrolysis of ptRNA^{Tyr} but not p4.5S RNA (Table 1; Figures 3 and 4). The observation that these mutants can help M1 RNA cleave ptRNA^{Tyr} indicates that these mutants do bind M1 RNA. However, the holoenzymes that are formed by the interaction of M1 RNA and these mutants (individually) must be different in their structure compared to the wild-type holoenzyme, as reflected by the narrower substrate specificity of the mutant holoenzymes. Kinetic analyses are needed to distinguish whether the inability of these mutant holoenzymes to cleave p4.5S RNA is due to weak substrate binding or to a slow rate of cleavage.

Function of mutant derivatives of C5 protein *in vivo*

There is a reasonable correlation between the ability of the various mutant derivatives to participate in RNase P catalysis *in vitro* and that observed *in vivo* (Table 1). At 43°C, mutants C5 F18A/W109A, C5 F22A/W109A, and C5 F18A/F73A are severely compromised in their ability to cleave ptRNA^{Tyr} and p4.5S RNA *in vitro* and are also unable to support growth of *T7A49* cells. C5 F22A and C5 R67A, whose ptRNA^{Tyr} cleavage activity is reduced 2.5- to 5-fold *in vitro*, exhibit only moderate complementation *in vivo*.

However, there are some mutants, such as C5 R62A, which catalyze the cleavage of ptRNA^{Tyr} quite efficiently but are able to support only moderate complementation *in vivo*. It is noteworthy that C5 R62A does not catalyze the hydrolysis of p4.5S RNA *in vitro*; this raises the possibility that a defect in 4.5S RNA biosynthesis might underlie the inability of these mutants to fully complement the ts phenotype. However, since our complementation assay examines the ability of *T7A49* cells to grow at 43°C, it is possible that heat shock proteins which are induced at 43°C are able to suppress the phenotypic effects of a defect in 4.5S RNA biosynthesis (Wood *et al.*, 1992). Therefore, any defect in 4.5S RNA biosynthesis (as a result of low RNase P activity) might not manifest as a severe growth defect in *T7A49* cells at 43°C.

Although an evaluation of *in vitro* and *in vivo* results leads to useful inferences, discrepancies in correlating *in vitro* with *in vivo* results may be the consequence of several factors. In this report, we have tested the activity of certain mutant derivatives of C5 protein *in vitro* with two of the substrates of RNase P. This might not be adequate to extrapolate and draw conclusions about the *in vivo* performance of these mutants considering that RNase P acts on as many as 60 different substrates *in vivo* and that the effects of C5 protein on RNase P catalysis are substrate-identity dependent (Kirsebom & Altman, 1989; Peck-Miller & Altman, 1991; Kirsebom & Svärd, 1992).

The failure of some of the mutant derivatives to complement the ts phenotype of *T7A49* cells could be due to lower affinity of the mutants, relative to wild-type C5 protein, for M1 RNA. It is to be expected then that overexpression of these mutants in *T7A49* cells would favor holoenzyme assembly and thus rescue the ts phenotype. However, the IPTG-induced overexpression of C5 wild-type protein (or its mutant derivatives) proved to be toxic to the cell regardless of growth temperature (data not shown).

Structural predictions of C5 protein

In the absence of a three-dimensional structure of C5 protein, it is not possible to explain fully the functional phenotypes observed with the mutant derivatives of C5 protein. Until X-ray crystallographic or NMR spectroscopic studies establish the tertiary structure of C5 protein, a working three-dimensional model of C5 protein will help to design rational mutants and dissect the role of various residues in the function of C5 protein.

An αββαβα core motif (Figure 5) was predicted for all the protein subunits of RNase P from bacteria by the PHD algorithm (Rost & Sander, 1993). Reported accuracy rates are 72% for prediction of protein secondary structure by this algorithm. Any secondary structure prediction is likely to be compatible with various tertiary structure models. However, we entertained the possibility that C5 protein could adopt a fold the core of which consists of an anti-parallel β-sheet flanked on one side by two α-helices, since this fold has already been observed in other RNA-binding proteins (for example, U1A, a spliceosomal protein and S6, a

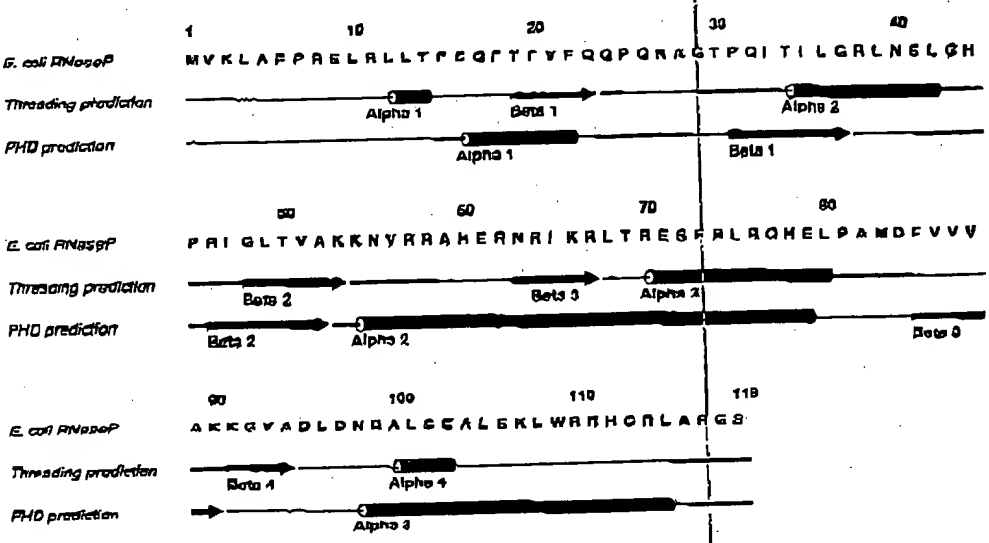


Figure 5. Predictions of the secondary structure of C5 protein. The predictions were obtained using the PHD algorithm (Rost & Sander, 1993) and based on threading analysis (Bryant & Lawrence, 1993) in which the sequences of C5 protein and its homologs were individually threaded through the X-ray structure of a human U1A mutant protein complexed with an RNA hairpin (Oubridge *et al.*, 1994). The threading analysis used only the protein coordinates from the B/Q complex in this PDB file. For further information on the data obtained from the threading analysis and for coordinates of the tertiary structure model, contact landsman@nih.gov.

ribosomal protein; Lindahl *et al.*, 1994; Nagai *et al.*, 1990, 1995; Nagai, 1996). Although the PHD prediction of αββαβα is different from the βαββαβα motif in U1A, we decided to use the "threading" technique to evaluate the possibility that C5 protein adopts a fold similar to that of U1A.

Threading (or homology model building) methods enable examination of the ability of a given query sequence to adopt a three-dimensional structure already in the protein database (Bryant & Lawrence, 1993) and are capable of revealing structural similarities which are not evident from conventional sequence alignments. Our computations indicate that despite the low degree of sequence identity among the 12 RNase P protein sequences from various bacteria, ten out of 12 sequences can adopt the U1A fold consisting of a four-stranded anti-parallel β-sheet with two α-helices packed on one side (unpublished observations). Based on the sequence-structure alignment generated using homology model building, we constructed a hypothetical three-dimensional model of C5 protein which identifies roles for various conserved residues. The β-sheet in C5 protein could provide a surface for docking the RNA ligand, as in other well-characterized RNA-binding domains (Oubridge *et al.*, 1994; Nagai *et al.*, 1995). A stretch of seven amino acid residues (RNRIKRL), containing four basic residues, is conserved in nearly all the RNase P protein subunits from bacteria. Interestingly, the NRIKR sequence is in strand β3 (according to the U1A nomenclature) and is the equivalent of the highly conserved RNP1 consensus motif in U1A which plays a crucial role in RNA binding. The surface electrostatic potential map of C5 protein reveals a binding cleft, rich in conserved, positively charged residues (such as Lys54, Arg62, and Lys66), which could anchor a loop present in M1 RNA (data not shown).

Although a number of correct structural predictions have been made with computational methods, it is critical to bear in mind that the secondary and tertiary structure predictions mentioned above serve only as working models of C5 protein. Further experimental data in support of or in contradiction of these predictions are required.

Materials and Methods

Materials

The various reagents used in this study were obtained from the indicated commercial sources: restriction enzymes, T4 DNA polymerase, T4 DNA ligase, T4 polynucleotide kinase, and *E. coli* DNA polymerase (Klenow fragment) were from New England Biolabs; T7 RNA polymerase, SP6 RNA polymerase, *Sma*I, RNasin and M13K07 helper phage were from Promega; nucleoside triphosphates and CS0 CM Sephadex were from Pharmacia Biotech; carbenicillin was from Gemini Bioproducts; BCA protein assay reagent was from Pierce; QuickSpin Sephadex columns were from Boehringer Mannheim; and [α-^{32}P]GTP was from Amersham Biochemicals.

Table 2. Oligonucleotides used for mutagenesis

Mutation(s)	Sequence of oligonucleotide used for site-directed mutagenesis	Name of plasmid
F18A	5′ C GAA AAC GAA GGT AGC CTG GGA CGG GG 3′	pBSC5Sn7
W109A	5′ G GCG GCG CGC TAA TTT TTC C 3′	pBSC5Sn8
R62A	5′ GT CAG ACG TTT GAT CCG ATT AGC TTC ATG GGC G 3′	pBSC5Sn9
N63I′	5′ GT CAG ACG TTT GAT CCG CGG GCG TTC ATG GG 3′	pBSC5Sn10
N63V	5′ GT CAG ACG TTT GAT CCG AAC CCC TTC ATG GC 3′	pBSC5Sn13
F22A	5′ C ACG CTG CGG CTG CTG CGC AAC GAA GG TGA ACT GGG 3′	pBSC5Sn16
F73A	5′ C ATC TTC CCC CAC ACC CGC GCT TTC ACG CGT CAC 3′	pBSC5Sn17
K66A	5′ G GCG CAG ACG GAA GCT TTC GCG AGT CAG ACG AGC AAT CCG ATT GCG TTC ATG GGC 3′	pBSC5Sn24
R67A	5′ G GCG CAG ACG GAA GCT TTC GCG AGT CAG AGC TTT AAT CCG ATT GCG TTC ATG GGC 3′	pBSC5Sn25
F18A, W109A	5′ G GAA AAC GAA GGT AGC CTG GGA CGG GG 3′ (F18A) 5′ G GCG GCG CGC TAA TTT TTC C 3′ (W109A)	pBSC5Sn78
F18W, W109F	5′ GAA AAC GAA CGT CCA CTG GGA CGG G 3′ (F18W) 5′ GTG GCG GCG GAA TAA TTT TTC C 3′ (W109F)	pBSC5Sn112
F18A, F22A	5′ C ACG CTG CGG CTG CTG CGC AAC GAA GGT AGC CTG GGA CGG GG 3′	pBSC5Sn716
F18A, F73A	5′ G GAA AAC GAA GGT AGC CTG GGA CGG GG 3′ (F18A) 5′ C ATG TTC GCC CAC ACC CGC GCT TTC ACG CGT CAG 3′ (F73A)	pBSC5Sn717
F22A, W109A	5′ C ACG CTG CGG CTG CTG CGC AAC GAA GG TGA ACT GGG 3′ (F22A) 5′ C CCC CCC CGC TAA TTT TTC C 3′ (W109A)	pBSC5Sn810
F73A, W109A	5′ C ATG TTC GCG CAG ACG CGC GCT TTC ACC CGT CAC 3′ (F73A) 5′ G GCG GCG CGC TAA TTT TTC C 3′ (W109A)	pBSC5Sn817

Since the single-stranded DNA used as the template for site-directed mutagenesis of C5 gene contains the gene in the sense orientation, the oligonucleotides used are in the antisense orientation. The bold letters indicate the codon that was altered. The underlined nucleotides represent changes in the wobble position (of the respective codon) that was introduced in addition to the desired mutation. These additional modifications (i.e. silent mutagenesis) led to changes in the restriction pattern of the mutant gene and enabled the rapid screening of mutants. In many cases there were no other changes in addition to the desired mutation since the mutation itself resulted in a restriction pattern that is distinct from that of the wild-type C5 gene.

Construction of pBSC5

Vlcque *et al.* (1988) had constructed the plasmid pARE7 in which a semisynthetic C5 gene was placed downstream from the promoter and the ribosome binding site of gene *10* of T7 bacteriophage. For objectives unrelated to this study, we constructed a plasmid pVG2, which is analogous to pARE7 except that it lacks the *EcoRI* site in the vector. In order to clone the C5 gene into a vector which possesses the f1 filamentous phage origin of replication, the entire nucleotide sequence containing the T7 RNA polymerase promoter, translational signals, the C5 coding region and the T7 transcription terminator was moved *en bloc* from pVG2 to pBluescript II KS (+) to generate the plasmid, pBSC5. This subcloning involved: (1) digesting pVG2 with *BglII* and *EcoRV*, (2) filling-in the 3′ recessed ends created by the *BglII* digest, and (3) ligating this blunt-ended DNA fragment into pBluescript II KS (+) which had been digested with *PvuII*. Upon digesting pBluescript II KS (+) with *PvuII*, the *lacI*, multiple cloning box, and the *lacZ* sequences present in between the two *PvuII* sites were deleted. In pBSC5, the C5 gene is cloned in the same orientation as *lacI* and *lacZ* in the parental vector (pBluescript II KS (+)).

Site-directed mutagenesis of C5 protein

Mutations were engineered in the gene encoding C5 protein using the "oligonucleotide-directed mutagenesis without phenotypic selection" procedure (Kunkel, 1985). Single-stranded (ss) DNA was prepared by coinfection of *CJ236* (pBSC5) with the helper phage M13K07. The ss DNA was isolated according to the manufacturer's instructions in the pALTER kit (Promega). Mutagenesis reactions were performed according to the protocol described by Kunkel (1989). The various mutant derivatives of C5 protein were generated using the DNA oligonucleotides described in Table 2. DNA oligonucleotides were synthesized at the Keck Biotechnology Resource Laboratory at Yale Medical School.

In the T7 promoter-driven system for overexpression of proteins (Studier *et al.*, 1990), it is preferable to clone the target gene initially into a host that does not contain the T7 RNA polymerase gene to ensure that there is no plasmid instability due to toxicity associated with expression of the target gene. Therefore, to obtain clones of the mutant derivatives of C5 protein, DH5α was transformed with the mutagenesis reactions and the plasmid DNA containing the desired mutation isolated. The presence of the engineered mutation was confirmed by sequencing of the various plasmid DNA samples. Subsequently, these plasmid DNAs were used to transform *BL21(DE3)* cells and the respective mutant proteins overexpressed.

Complementation assay *in vivo*

T7A49 cells containing both the *rnpA49* mutation (i.e. C5 R46H) in their chromosomes and the T7 RNA polymerase gene under control of the *lac UV5* promoter (Guerrier-Takada *et al.*, 1995) were transformed with plasmids bearing the wild-type C5 protein or its mutant derivatives under the control of a promoter for T7 RNA polymerase transcription. The permissive temperature for this strain is 30 C while the non-permissive tempera-

ture is 43 C. The transformants bearing the various mutants were then grown overnight at 30 C in LB media supplemented with carbenicillin (100 μg/ml). Cultures were reinoculated the next morning in fresh LB carbenicillin media and grown at 30 C until the A_{600} ~ 0.20. The cultures were then shifted to 43 C and the cell growth at the non-permissive temperature monitored by measuring A_{600} at regular time intervals. Cells (without any plasmids) were grown in the absence of carbenicillin and served as the negative control while cells transformed with pBSC5 (wild-type C5 protein) served as the positive control for complementation of the ts phenotype.

Overexpression and purification of mutant derivatives of C5 protein

BL21(DE3) cells containing plasmids encoding the various mutant derivatives of C5 protein (see Table 2) were grown to A_{600} ~ 0.4 and then induced with 2 mM IPTG. After establishing in small scale cultures (approximately 4 ml) that there was T7 promoter-driven overexpression of the various proteins, large scale (1 *l*) cultures were grown to isolate the respective proteins. When the crude cell extract was prepared either by sonication or by grinding the cells with alumina and subsequently centrifuged at 30,000*g* for 30 minutes, it was discovered that the mutant proteins were present in the P30 (i.e. the pellet obtained after the centrifugation) rather than the S30 (i.e the supernatant obtained after the centrifugation). The relative amounts in the P30 and the S30 varied for the various mutants, perhaps indicative of the folding properties of the various mutants. The mutant proteins were isolated from the P30 following the protocol described by Baer *et al.* (1989). This purification scheme for the isolation of C5 R46H involves solubilizing the P30 with 4 M urea (in 50 mM Tris-hydrochloride (pH 7.5), 10 mM magnesium acetate, 60 mM ammonium chloride) in order to recover the protein (Baer *et al.*, 1989). The urea-solubilized suspension was then centrifuged for 30 minutes at 50,000*g*. The P30 wash was dialyzed against 50 mM Tris-hydrochloride (pH 7.5), 100 mM ammonium chloride, 10 mM magnesium chloride, 10 mM dithiothreitol (DTT). The protein precipitates upon removal of urea. The precipitated protein is then resuspended in a buffer containing 7 M urea and further purified using CM Sephadex C50 chromatography. Our modifications from this procedure are: (1) the cells were lysed using a sonicator instead of being crushed in alumina, and (2) the deliberate omission of the reducing agent before loading the protein on a CM Sephadex column. In the absence of DTT, dimer formation is promoted. The dimeric version of the C5 mutants elutes at a higher salt concentration, relative to the monomer, and can help purify the C5 mutant proteins from contaminants that elute at lower salt concentrations.

Aliquots from the various fractions eluted from CM Sephadex columns were electrophoresed on SDS-polyacrylamide gels and stained with silver nitrate. Only those fractions which showed high purity (>95%) were pooled and the stocks were stored at −70 C. Immediately before use in RNase P assays, the various mutant preparations were thawed and treated with 10 mM DTT. The protein concentration of the various preparations was assessed using the BCA test. The standard curve for the BCA test was generated using wild-type C5 protein, the concentration of which was determined by measuring Trp absorbance at 280 nm.

Assays for RNase P activity *in vitro*

Plasmid encoding M1 RNA (pJA2') was linearized with *Fok*I and transcribed by T7 RNA polymerase as described by Vioque *et al.* (1988). The RNA was then isolated using a QuickSpin column procedure (Vioque *et al.*, 1988). Plasmids encoding ptRNA^{Tyr} and p4.5S RNA were linearized with *Fok*I and *Sma*I, respectively, and these RNA substrates were internally labeled with [α-^{32}P]GTP during *in vitro* transcription with T7 RNA polymerase and then purified on denaturing 8% (w/v) polyacrylamide/7 M urea gels.

Cleavage of RNA substrates was performed with RNase P holoenzymes reconstituted using 1 nM M1 RNA and 200 nM of C5 protein or its mutant derivatives. The assays were performed with 100 nM ptRNA^{Tyr} or p4.5S RNA at 30 C or 43 C. The cleavage reactions were performed in 10 mM Hepes (pH 7.5), 10 mM magnesium acetate, 400 mM ammonium acetate, 5% (v/v) glycerol and 0.01% (v/v) NP-40. The ptRNA^{Tyr} assays were carried out at 30 C and 43 C for ten minutes and five minutes, respectively. The p4.5S RNA assays were carried out at 30 C and 43 C for 20 minutes and ten minutes, respectively. Autoradiograms were obtained after separating the products of the various reactions on an 8% (for ptRNA^{Tyr}) or 7% (for p4.5S RNA) polyacrylamide/7 M urea gel. The autoradiograms were scanned using a Microtek MSF-300Z scanner and the Scan Maker Plug-in software for Adobe Photoshop version 3.0. In order to group the hydrophobic mutants together, the autoradiograms from two different experiments were used to generate the composites shown in Figures 3 and 4.

To calculate the initial velocity, the enzyme assays were performed at the indicated temperature and aliquots withdrawn at regular time intervals and quenched with 9 M urea. The products of the reaction were electrophoresed on polyacrylamide/7 M urea gels and the extent of cleavage was calculated by quantifying the intensity of the various bands using a phosphorimager (Fuji). The extent of cleavage was always restricted to the linear range for product formation.

Database searches

The GenPept release 91.0 (Benson *et al.*, 1996), EMBL release 43.0 (Rodriguez-Tome *et al.*, 1996), PIR version 45.0 (George *et al.*, 1996), and Swiss-Prot version 31.0 (Bairoch & Apweiler, 1996) databases were searched using the BLASTP algorithm (Altschul *et al.*, 1990), with the *E. coli* RNase P sequences used as the basis for comparison. BLAST search cutoffs used to identify homologs were a Karlin/Altschul score for two aligned sequence segments >70 with a probability of $<10^{-1}$. These database searches identified 12 full-length RNase P sequences.

Acknowledgements

We are grateful to members of our laboratory, especially Drs Cecilia Guerrier-Takada, Paul Eder and Urban Lundberg, for helpful discussions. Research in the laboratory of S.A. is supported by National Institutes of Health grant GM19422. V.G. was funded by postdoctoral fellowships from the Patrick and Catherine Donaghue Foundation for Medical Research and the Anna Fuller Fund for Cancer Research.

References

Alifano, P., Rivellini, F., Piscitelli, C., Arraiano, C. M., Bruni, C. B. & Carlomagno, M. S (1994). RNase E provides substrates for ribonuclease P-dependent processing of polycistronic RNA. *Genes Dev.* **8**, 3021–3031.

Altman, S., Kirsebom, L. & Talbot, S. J. (1995). Recent studies of RNase P. In *tRNA: Structure, Biosynthesis and Function* (Soll, D. & RajBhandary, U., eds), pp. 67–78. American Society for Microbiology, Washington, DC.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.

Baer, M. F., Wesolowski, D. & Altman, S. (1989). Characterization *in vitro* of the defect in a temperature-sensitive mutant of the protein subunit of RNase P from *Escherichia coli*. *J. Bacteriol.* **171**, 6862–6866.

Bairoch, A. & Apweiler, R. (1996). The SWISS-PROT protein sequence data bank and its new supplement TREMBL. *Nucl. Acids Res.* **24**, 21–25

Benson, D. A., Boguski, M., Lipman, D. J. & Ostell, J. (1996). GenBank. *Nucl. Acids Res.* **24**, 1–5.

Bothwell, A. L. M., Garber, R. L. & Altman, S. (1976). Nucleotide sequence and *in vitro* processing of a precursor molecule to Escherichia coli 4.5S RNA. *J. Biol. Chem.* **251**, 7709–7716.

Bryant, S. H. & Lawrence, C. E. (1993). An empirical energy function for threading protein sequence through a folding motif. *Proteins: Struct. Funct. Genet.* **16**, 92–113.

Burd, C. G. & Dreyfuss, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins. *Science*, **265**, 615–621.

Cordes, M. H. J., Davidson, A. R. & Sauer, R. T. (1996). Sequence space, folding and protein design. *Curr. Opin. Struct. Biol.* **6**, 3–10.

George, D. G., Barker, W. C., Mewes, H. W., Pfeiffer, F. & Tsugita, A. (1996). The PIR-international protein sequence database. *Nucl. Acids Res.* **24**, 17–20.

Gopalan, V., Golbik, R., Schreiber, G., Fersht, A. R. & Altman, S. (1997). Fluorescence properties of a tryptophan residue in an aromatic core of the protein subunit of ribonuclease P from *Escherichia coli*. *J. Mol. Biol.* **267**, 765–769.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. R. & Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, **35**, 849–857.

Guerrier-Takada, C., Li, Y. & Altman, S. (1995). Artificial regulation of gene expression in *Escherichia coli* by RNase P. *Proc. Natl Acad. Sci. USA*, **92**, 11115–11119.

Kirsebom, L. A. & Altman, S. (1989). Reaction *in vitro* of some mutants of RNase P with wild-type and temperature-sensitive substrates. *J. Mol. Biol.* **207**, 837–840.

Kirsebom, L. A. & Svärd, S. G. (1992) The kinetics and specificity of cleavage by RNase P is mainly dependent on the structure of the amino acid acceptor stem. *Nucl. Acids Res.* **20**, 425–432.

Kirsebom, L. A., Baer, M. F. & Altman, S. (1988). Differential effects of the mutations in the protein and RNA moieties of RNase P on the efficiency of suppression by various tRNA suppressors. *J. Mol. Biol.* **204**, 879–888.

Komine, Y., Kitabatake, M., Yokogawa, T., Nishikawa, K. & Inokuchi, H. (1994). A tRNA-like structure is present in 10Sa RNA, a small stable RNA from *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **91**, 9223–9227.

Kunkel, T. A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc Natl Acad. Sci. USA*, **82**, 488–492.

Kunkel, T. A. (1989). Mutagenesis of cloned DNA. In *Current Protocols in Molecular Biology* (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K., eds), pp. 8.1.1–8.1.6. John Wiley, New York.

Laziński, D., Grzadzielska, E. & Das, A. (1989). Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. *Cell*, **59**, 207–218.

Lindahl, M., Svensson, L. A., Liljas, A., Sedelnikova, S. E., Eliseikina, I. A., Fomenkova, N. P., Nevskaya, N., Nikonov, S. V., Garber, M. B., Muranova, T. A., Rykonova, A. I. & Amons, R. (1994). Crystal structure of the ribosomal protein S6 from *Thermus thermophilus*. *EMBO J.* **13**, 1249–1254.

Lumelsky, N. & Altman, S. (1988). Selection and characterization of randomly produced mutants in the gene coding for M1 RNA. *J. Mol. Biol.* **202**, 443–454.

Matthews, B. W. (1993). Structural and genetic analysis of protein folding and stability. *Curr. Opin. Struct. Biol.* **3**, 589–593.

Nagai, K. (1996). RNA-protein complexes. *Curr. Opin. Struct. Biol.* **6**, 53–61.

Nagai, K., Oubridge, C., Jessen, T. H., Li, J. & Evans, P. R. (1990). Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature*, **348**, 515–520.

Nagai, K., Oubridge, C., Ito, N., Avis, J. & Evans, P. R. (1995). The RNP domain: a sequence-specific RNA-binding domain involved in processing and transport of RNA. *Trends Biochem. Sci.* **20**, 235–240.

Oubridge, C., Ito, N., Evans, P. R., Teo, C. H. & Nagai, K. (1994). Crystal structure at 1.92 Å resolution of the RNA-binding domain of U1A spliceosomal protein complexed with an RNA hairpin. *Nature*, **372**, 432–438.

Pace, N. R. & Brown, J. W. (1995). Evolutionary perspective on the structure and function of ribonuclease P, a ribozyme. *J. Bacteriol.* **177**, 1919–1928.

Peck-Miller, K. & Altman, S. (1991). Kinetics of the processing of the precursor to 4.5S RNA, a naturally occurring substrate for RNase P from *Escherichia coli*. *J. Mol. Biol.* **221**, 1–5.

Reich, C., Olsen, G. J., Pace, B. & Pace, N. R. (1988). Role of the protein moiety of RNase P, a ribonucleoprotein enzyme. *Science*, **239**, 179–182.

Rodrigues-Tome, P., Stoehr, P. J., Cameron, G. N. & Flores, T. P. (1996). The European Bioinformatics Institute (EBI) databases. *Nucl. Acids Res.* **24**, 6–12.

Rost, B. & Sander, C. (1993). Prediction of protein secondary structure at better than 70 accuracy *J. Mol Biol.* **232**, 584–596.

Schedl, P. & Primakoff, P. (1973). Mutants of *Escherichia coli* thermosensitive for the synthesis of transfer RNA. *Proc. Natl Acad. Sci. USA*, **70**, 2091–2095.

Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dunbendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89.

Talbot, S. J. & Altman, S (1994). Kinetic and thermodynamic analysis of RNA-protein interactions in the

ribonuclease P holoenzyme from *Escherichia coli*. *Biochemistry*, 33, 1406–1412.

Thompson, J. D., Higgins, D. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* 22, 4673–4680.

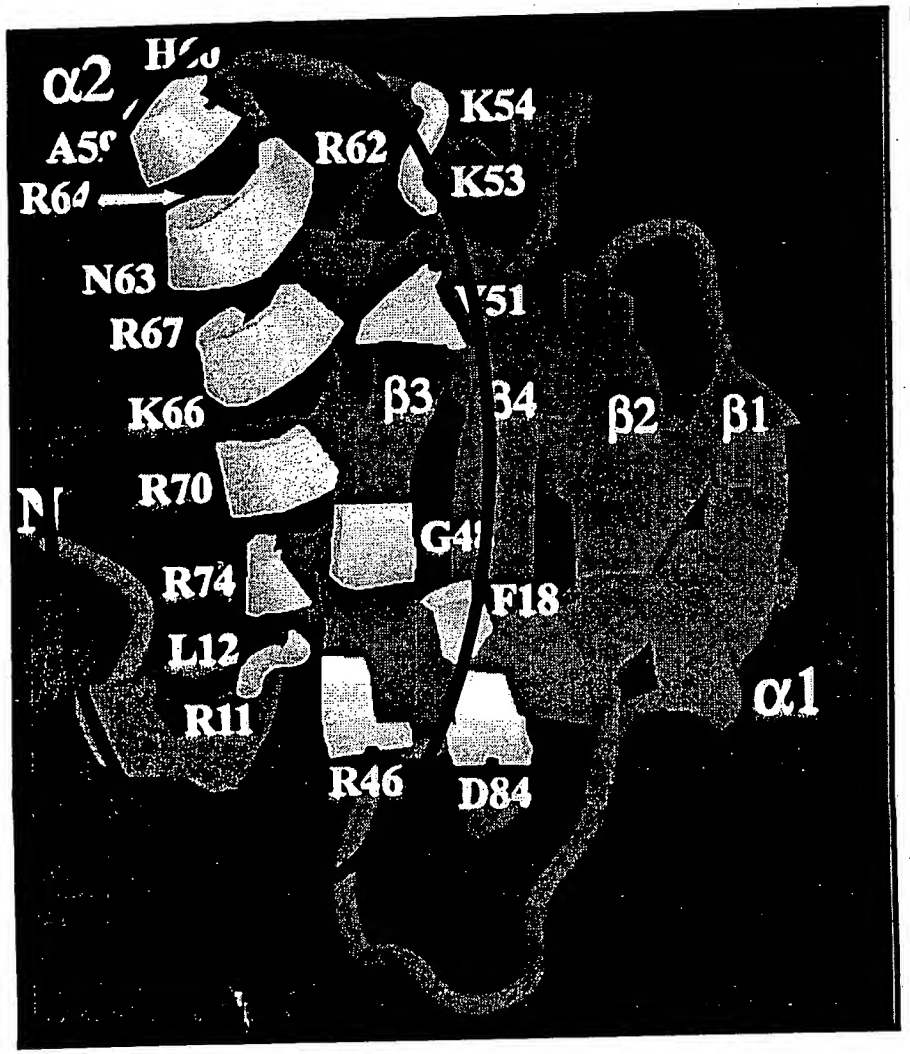
Vioque, A., Arnez, J. & Altman, S. (1988). Protein-RNA interactions in the RNase P holoenzyme from *Escherichia coli*. *J. Mol. Biol.* 202, 835–848.

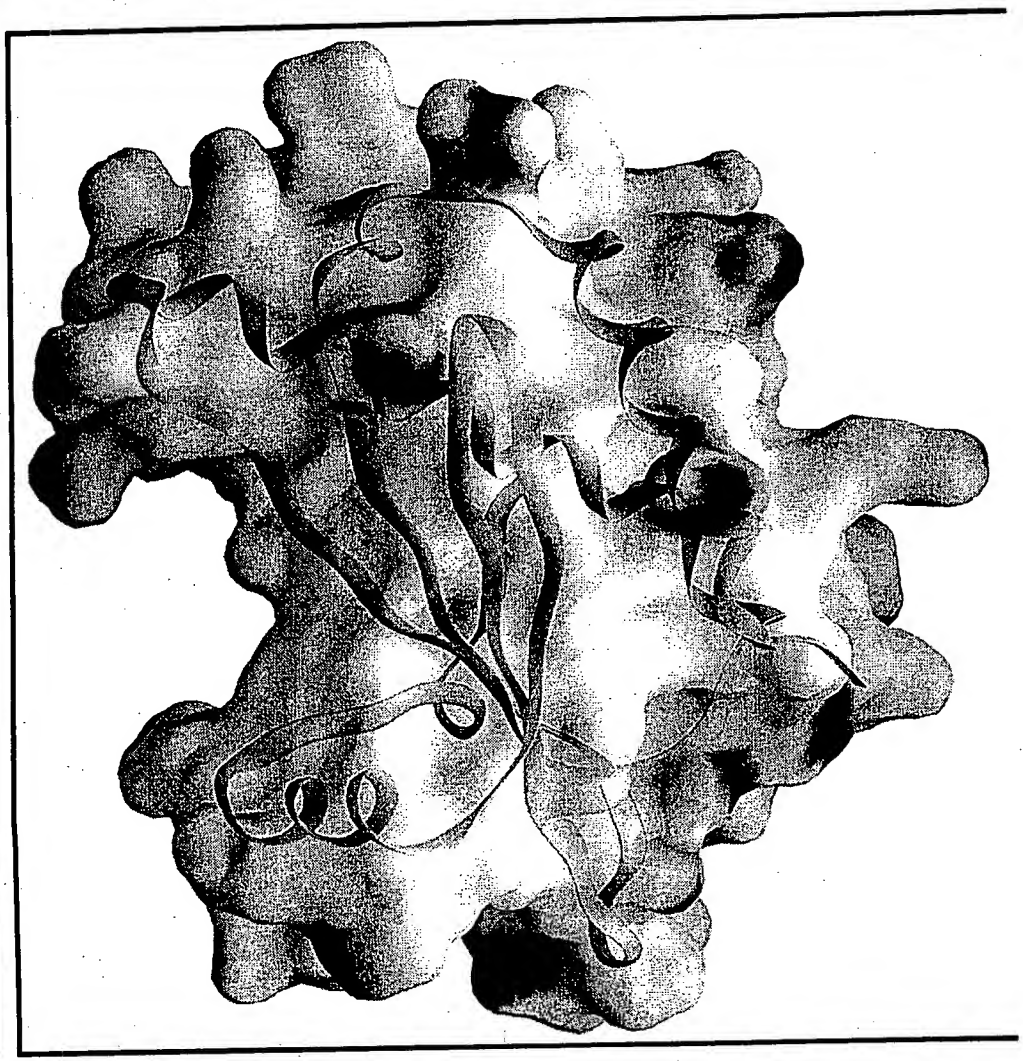
Wood, H., Lorink, J. & Tollervey, D. (1992). Evolutionarily conserved nucleotides within *E. coli* 4.5S RNA required for association with P48 *in vitro* and for optimal function *in vivo*. *Nucl. Acids Res.* 20, 5919–5925.

Edited by K. Nagai

(Received 21 August 1996; received in revised form 13 December 1996; accepted 6 January 1997)

Location of conserved residues in the tertiary structure of bacterial RNase P





Comparative Genomics Tools

Keyword Search: [help]

COG594

Select Genome(s):

-- Favorites
Escherichia coli K12
Bacillus subtilis
Shewanella oneidensis MR-1
Desulfuromonas acetoxidans
Desulfovibrio vulgaris
Desulfovibrio desulfuricans G20
Geobacter metallireducens
Geobacter sulfurreducens PCA
Bdellovibrio bacteriovorus HD100
Desulfotalea psychrophila
-- Bacteria
Aquifex aeolicus VF5
Thermotoga maritima
Deinococcus radiodurans
Chlorobium tepidum TLS
Bacteroides thetaiotaomicron VPI-5482
Fusobacterium nucleatum, ATCC25586
---- Cyanobacteria
Thermosynechococcus elongatus BP-1

Search | Info | GO | Pathways

Search results for COG594 in 131 genomes

From COG: 105 found.

	Genes	
1	**G O D H S B** VIMSS802 : **rnpA** CT784, NCBI ptt file:Ribonuclease P Protein Component **COG594**, RNase P protein component	*Chlamydia trachomatis*
2	**G O D H S B** VIMSS2341 : TM1463, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Thermotoga maritin*
3	**G O D H S B** VIMSS7350 : **HI0999** HI0999, NCBI ptt file:ribonuclease P (rnpA) **COG594**, RNase P protein component	*Haemophilus influenzae Rd KW2*
4	**G O D H S B** VIMSS8554 : **rnpA** MG465, NCBI ptt file:ribonuclease P protein component (rnpA) **COG594**, RNase P protein component	*Mycoplasma genitalium*
5	**G O D H S B** VIMSS11953 : **rnpA** slr1469, NCBI ptt file:protein subunit of ribonuclease P **COG594**, RNase P protein component	*Synechocystis sp. PCC 6803*
6	**G O D H S B** VIMSS14138 : **rnpA** MPN681, NCBI ptt file:RNaseP C5 chain **COG594**, RNase P protein component	*Mycoplasma pneumoniae*

7	**G O D H S B** VIMSS17763 : **rnpA** b3704, NCBI ptt file:RNase P, protein component; protein C5; processes tRNA, 4.5S RN **COG594**, RNase P protein component	*Escherichia coli K1:*
8	**G O D H S B** VIMSS19859 : HP1448, NCBI ptt file:ribonuclease P, protein component (rnpA) **COG594**, RNase P protein component	*Helicobacter pylori 26695*
9	**G O D H S B** VIMSS26767 : TP0950a, NCBI ptt file:RNAse P, protein component **COG594**, RNase P protein component	*Treponema pallidur.*
10	**G O D H S B** VIMSS28195 : **rnpA** jhp1341, NCBI ptt file:putative RIBONUCLEASE P PROTEIN COMPONENT **COG594**, RNase P protein component	*Helicobacter pylori J99*
11	**G O D H S B** VIMSS29265 : **rnpA** CPn0934, NCBI ptt file:Ribonuclease P Protein Component **COG594**, RNase P protein component	*Chlamydophila pneumoniae CWL0.*
12	**G O D H S B** VIMSS35697 : **rnpA** Rv3923c, NCBI ptt file:rnpA **COG594**, RNase P protein component	*Mycobacterium tuberculosis H37Rv*
13	**G O D H S B** VIMSS36281 : RP611, NCBI ptt file:RIBONUCLEASE P (rnpA) **COG594**, RNase P protein component	*Rickettsia prowazel*
14	**G O D H S B** VIMSS40644 : **rnpA** Bsu4102, NCBI ptt file:ribonuclease P (protein component) **COG594**, RNase P protein component	*Bacillus subtilis*
15	**G O D H S B** VIMSS42793 : **DR2151** DR2151,	*Deinococcus radiodurans*

	NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	
16	**G O D H S B** VIMSS44083 : BB0441, NCBI ptt file:ribonuclease P protein component (rnpA) **COG594**, RNase P protein component	*Borrelia burgdorferi B31*
17	**G O D H S B** VIMSS45950 : **rnpA** UU603, NCBI ptt file:ribonuclease p - protein component **COG594**, RNase P protein component	*Ureaplasma urealyticum*
18	**G O D H S B** VIMSS46841 : **rnpA** Cj0960c, NCBI ptt file:putative ribonuclease P protein component **COG594**, RNase P protein component	*Campylobacter jeju*
19	**G O D H S B** VIMSS48496 : CP0927, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Chlamydophila pneumoniae AR39*
20	**G O D H S B** VIMSS51479 : XF2781, NCBI ptt file:ribonuclease P **COG594**, RNase P protein component	*Xylella fastidiosa 9a*
21	**G O D H S B** VIMSS52478 : **rnpA** CPj0934, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Chlamydophila pneumoniae J138*
22	**G O D H S B** VIMSS52621 : VC0006, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Vibrio cholerae*
23	**G O D H S B** VIMSS62016 : **rn** PA5569, NCBI ptt file:ribonuclease P protein component	*Pseudomonas aeruginosa PAO1*

COG594, RNase P protein component

24 **G O D H S B** VIMSS62031 : **rnpA** BU014, *Buchnera aphidicol str. APS (Acyrthosiphon pisu*

NCBI ptt file:ribonuclease P protein component
COG594, RNase P protein component

25 **G O D H S B** VIMSS66646 : **rnpA** BH4065, *Bacillus halodurans*

NCBI ptt file:ribonuclease P
COG594, RNase P protein component

26 **G O D H S B** VIMSS76129 : **rnpA** Z5195, *Escherichia coli O157:H7 EDL933*

NCBI ptt file:RNase P, protein component; protein C5; processes tRNA, 4.5S RN
COG594, RNase P protein component

27 **G O D H S B** VIMSS76936 : **rnpA** L131443, *Lactococcus lactis subsp. lactis*

NCBI ptt file:ribonuclease P protein component (EC 3.1.26.5)
COG594, RNase P protein component

28 **G O D H S B** VIMSS80238 : **rnpA** PM1163, *Pasteurella multoci*

NCBI ptt file:RnpA
COG594, RNase P protein component

29 **G O D H S B** VIMSS82694 : **rnpA** ML2712, *Mycobacterium lepr*

NCBI ptt file:ribonuclease P protein component
COG594, RNase P protein component

30 **G O D H S B** VIMSS86395 : mlr4810, *Mesorhizobium loti*

NCBI ptt file:ribonuclease P (protein component)
COG594, RNase P protein component

31 **G O D H S B** VIMSS96108 : ECs4639, *Escherichia coli O157:H7*

NCBI ptt file:ribonuclease P protein component
COG594, RNase P protein component

32	**G O D H S B** VIMSS97595 : CC0768, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Caulobacter crescentus CB15*
33	**G O D H S B** VIMSS100746 : **rnpA** SPy0246, NCBI ptt file:putative ribonuclease P protein component **COG594**, RNase P protein component	*Streptococcus pyogenes M1 GAS*
34	**G O D H S B** VIMSS104857 : **rnpA** SA2502, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Staphylococcus aureus subsp. aure N315*
35	**G O D H S B** VIMSS112020 : MT4041, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Mycobacterium tuberculosis CDC1*
36	**G O D H S B** VIMSS114735 : **rnpA** SAV2713, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Staphylococcus aureus subsp. aure Mu50*
37	**G O D H S B** VIMSS114889 : MYPU_1530, NCBI ptt file:RIBONUCLEASE P PROTEIN COMPONENT (PROTEIN C5) (RNASE P) **COG594**, RNase P protein component	*Mycoplasma pulmo*
38	**G O D H S B** VIMSS117466 : SP2042, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Streptococcus pneumoniae TIGR4*
39		

	G O D H S B VIMSS121317 : **rnpA** CAC3738, NCBI ptt file:RnpA, ribonuclease P protein component **COG594**, RNase P protein component	*Clostridium acetobutylicum*
40	**G O D H S B** VIMSS123120 : **rnpA** SMc01720, NCBI ptt file:PROBABLE RIBONUCLEASE P PROTEIN COMPONENT (PROTE C5) **COG594**, RNase P protein component	*Sinorhizobium melil*
41	**G O D H S B** VIMSS134736 : **rnpA** spr1853, NCBI ptt file:Ribonuclease P - protein component **COG594**, RNase P protein component	*Streptococcus pneumoniae R6*
42	**G O D H S B** VIMSS135865 : **rnpA** RC0937, NCBI ptt file:ribonuclease P [EC:3.1.26.5] **COG594**, RNase P protein component	*Rickettsia conorii*
43	**G O D H S B** VIMSS140961 : NMB1905, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Neisseria meningiti MC58*
44	**G O D H S B** VIMSS141717 : **rnpA** NMA0550, NCBI ptt file:putative ribonuclease P protein component **COG594**, RNase P protein component	*Neisseria meningiti Z2491*
45	**G O D H S B** VIMSS147357 : **rnpA** YPO4101, NCBI ptt file:ribonuclease P protein **COG594**, RNase P protein component	*Yersinia pestis CO*
46	**G O D H S B** VIMSS151121 : **rnpA** STM3840, NCBI ptt file:RNase P, protein component (protein C5), processes tRNA, 4.5S R **COG594**, RNase P protein component	*Salmonella typhimurium LT2*

47	**G O D H S B** VIMSS155329 : **rnpA** STY3939, NCBI ptt file:RNase P, protein component **COG594**, RNase P protein component	*Salmonella enterica subsp. enterica serovar Typhi*
48	**G O D H S B** VIMSS159076 : **rnpA** lmo2855, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Listeria monocytogenes EG e*
49	**G O D H S B** VIMSS162044 : **rnpA** lin2987, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Listeria innocua*
50	**G O D H S B** VIMSS165766 : alr3413, NCBI ptt file:ribonuclease P **COG594**, RNase P protein component	*Nostoc sp. PCC 71:*
51	**G O D H S B** VIMSS168279 : **rnpA** RSc0002, NCBI ptt file:PROBABLE RIBONUCLEASE P PROTEIN COMPONENT **COG594**, RNase P protein component	*Ralstonia solanacearum*
52	**G O D H S B** VIMSS173769 : **rnpA** Atu0385, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Agrobacterium tumefaciens str. C5 (U. Washington)*
53	**G O D H S B** VIMSS181130 : BMEII0276, NCBI ptt file:RIBONUCLEASE P PROTEIN COMPONENT **COG594**, RNase P protein component	*Brucella melitensis*
54	**G O D H S B** VIMSS187257 : **rnpA** CPE2659,	*Clostridium perfringens*

	NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	
55	**G O D H S B** VIMSS193405 : FN0002, NCBI ptt file:Ribonuclease P protein component **COG594**, RNase P protein component	*Fusobacterium nucleatum subsp. nucleatum ATCC 25586*
56	**G O D H S B** VIMSS195055 : **rnpA** spyM18_0228, NCBI ptt file:putative ribonuclease P protein component **COG594**, RNase P protein component	*Streptococcus pyogenes MGAS82*
57	**G O D H S B** VIMSS197322 : **tll0601** , NCBI ptt file:ORF_ID:tll0601~probable ribonuclease P **COG594**, RNase P protein component	*Thermosynechococc elongatus BP-1*
58	**G O D H S B** VIMSS199203 : **rnpA** SO0006, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Shewanella oneidensis MR-1*
59	**G O D H S B** VIMSS213458 : CT0004, NCBI ptt file:ribonuclease P protein component, putative **COG594**, RNase P protein component	*Chlorobium tepidun TLS*
60	**G O D H S B** VIMSS215714 : **rnpA** PP0008, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Pseudomonas putic KT2440*
61	**G O D H S B** VIMSS222895 : **rnpA** CBU1918,	*Coxiella burnetii RS 493*

	NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	
62	**G O D H S B** VIMSS227382 : **rnpA** SAV4314, NCBI ptt file:putative RNase P component **COG594**, RNase P protein component	*Streptomyces avermitilis MA-468(*
63	**G O D H S B** VIMSS231463 : **rnpA** CCA00835, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Chlamydophila cavi GPIC*
64	**G O D H S B** VIMSS240452 : **RnpA** TTE2801, NCBI ptt file:RNase P protein component **COG594**, RNase P protein component	*Thermoanaerobacte tengcongensis*
65	**G O D H S B** VIMSS244159 : **rnpA** SCO3881, NCBI ptt file:putative ribonuclease P component **COG594**, RNase P protein component	*Streptomyces coelicolor A3(2)*
66	**G O D H S B** VIMSS255516 : **rnpA** XCC4241, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Xanthomonas campestris pv. campestris str. ATC 33913*
67	**G O D H S B** VIMSS260213 : **rnpA** XAC4373, NCBI ptt file:ribonuclease P, protein component **COG594**, RNase P protein component	*Xanthomonas axonopodis pv. citri str. 306*
68	**G O D H S B** VIMSS262845 : **rnpA** ,	*Staphylococcus aureus subsp. aure*

		MW2
	NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	
69	**G O D H S B** VIMSS268156 : **rnpA** BA5737, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Bacillus anthracis s Ames*
70	**G O D H S B** VIMSS268171 : **rnpA** BUsg014, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Buchnera aphidicol str. Sg (Schizaphis graminum)*
71	**G O D H S B** VIMSS268877 : **rnpA** SpyM3_0175, NCBI ptt file:putative ribonuclease P protein component **COG594**, RNase P protein component	*Streptococcus pyogenes MGAS31*
72	**G O D H S B** VIMSS274628 : **rnpA** y4115, NCBI ptt file:RNase P **COG594**, RNase P protein component	*Yersinia pestis KIM*
73	**G O D H S B** VIMSS275063 : **rnpA** SAG0408, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Streptococcus agalactiae 2603V/R*
74	**G O D H S B** VIMSS280280 : **rnpA** OB3495, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Oceanobacillus iheyensis*
75	**G O D H S B** VIMSS280878 : **rnpA** BL0642a, NCBI ptt file:ribonuclease P protein component	*Bifidobacterium longum NCC2705*

COG594, RNase P protein component

76 **G O D H S B** VIMSS285104 : **rnpA** BRA1022, *Brucella suis 1330*
NCBI ptt file:ribonuclease P, protein component
COG594, RNase P protein component

77 **G O D H S B** VIMSS288810 : **rnpA** SF3760, *Shigella flexneri 2a str. 301*
NCBI ptt file:RNase P, protein component; protein C5; processes tRNA, 4.5S RN
COG594, RNase P protein component

78 **G O D H S B** VIMSS294305 : Wbr0124, *Wigglesworthia glossinidia endosymbiont of Glossina brevipalpi*
NCBI ptt file:RNase P protein component
COG594, RNase P protein component

79 **G O D H S B** VIMSS295137 : **rnpA** SMU.336, *Streptococcus mute UA159*
NCBI ptt file:putative ribonuclease P protein component
COG594, RNase P protein component

80 **G O D H S B** VIMSS297236 : gbs0443, *Streptococcus agalactiae NEM316*
NCBI ptt file:Unknown
COG594, RNase P protein component

81 **G O D H S B** VIMSS301835 : CE2946, *Corynebacterium efficiens YS-314*
NCBI ptt file:putative ribonuclease P protein component
COG594, RNase P protein component

82 **G O D H S B** VIMSS306367 : **rnpA** c4628, *Escherichia coli CFT073*
NCBI ptt file:Ribonuclease P protein component
COG594, RNase P protein component

83	**G O D H S B** VIMSS308254 : MYPE10410, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Mycoplasma penetrans*
84	**G O D H S B** VIMSS309204 : VV11005, NCBI ptt file:RNase P protein component **COG594**, RNase P protein component	*Vibrio vulnificus CMCP6*
85	**G O D H S B** VIMSS315210 : SE2418, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Staphylococcus epidermidis ATCC 12228*
86	**G O D H S B** VIMSS323308 : **rnpA** blr8097, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Bradyrhizobium japonicum*
87	**G O D H S B** VIMSS323542 : **rnpA** bbp014, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Buchnera aphidicol str. Bp (Baizongia pistaciae)*
88	**G O D H S B** VIMSS324815 : **rnpA** TW817, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Tropheryma whippl TW08/27*
89	**G O D H S B** VIMSS326850 : **rnpA** PD2122, NCBI ptt file:ribonuclease P **COG594**, RNase P protein component	*Xylella fastidiosa Temecula1*
90		

	G O D H S B VIMSS333040 : TW807, NCBI ptt file:unknown **COG594**, RNase P protein component	*Tropheryma whippl str. Twist*
91	**G O D H S B** VIMSS338511 : **rnpA** PSPTO5614, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Pseudomonas syringae pv. tomato str. DC3000*
92	**G O D H S B** VIMSS338516 : VP0004, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Vibrio parahaemolyticus RIMD 2210633*
93	**G O D H S B** VIMSS343524 : SPs0180, NCBI ptt file:putative ribonuclease P protein component **COG594**, RNase P protein component	*Streptococcus pyogenes SSI-1*
94	**G O D H S B** VIMSS348635 : **rnpA** t3679, NCBI ptt file:ribonuclease P **COG594**, RNase P protein component	*Salmonella enterica subsp. enterica serovar Typhi Ty2*
95	**G O D H S B** VIMSS352754 : BT3227, NCBI ptt file:putative ribonuclease P protein component **COG594**, RNase P protein component	*Bacteroides thetaiotaomicron VI 5482*
96	**G O D H S B** VIMSS357418 : **rnpA** EF3332, NCBI ptt file:ribonuclease P protein component	*Enterococcus faeca V583*

	COG594, RNase P protein component	
97	**G O D H S B** VIMSS362680 : BC5489, NCBI ptt file:Ribonuclease P protein component **COG594**, RNase P protein component	*Bacillus cereus ATC 14579*
98	**G O D H S B** VIMSS366122 : **rnpA** S4011, NCBI ptt file:RNase P, protein component; protein C5 **COG594**, RNase P protein component	*Shigella flexneri 2a str. 2457T*
99	**G O D H S B** VIMSS367124 : **rnpA** NE0389, NCBI ptt file:Bacterial ribonuclease P protein **COG594**, RNase P protein component	*Nitrosomonas europaea ATCC 19718*
100	**G O D H S B** VIMSS369218 : **rnpA** MGA_0630, NCBI ptt file:RnpA **COG594**, RNase P protein component	*Mycoplasma gallisepticum R*
101	**G O D H S B** VIMSS370486 : **rnpA** TC0167, NCBI ptt file:ribonuclease P protein component **COG594**, RNase P protein component	*Chlamydia muridarı*
102	**G O D H S B** VIMSS377048 : NCgl2992, NCBI ptt file:RNase P protein component **COG594**, RNase P protein component	*Corynebacterium glutamicum ATCC 13032*
103	**G O D H S B** VIMSS394464 : **rnpA** , VIMSS-AUTO:Ribonuclease P protein component (rnpA) **COG594**, RNase P protein component	*Desulfovibrio desulfuricans G20*

104	**G O D H S B** VIMSS206510 : **rnpA** DVU1075,ORF00727 TIGR:ribonuclease P protein component **COG594**, RNase P protein component	*Desulfovibrio vulgai*
105	**G O D H S B** VIMSS414059 : **rnpA** Bd3913, **COG594**, RNase P protein component	*Bdellovibrio bacteriovorus HD1(*

A description of the putative bacterial RNase P protein subunit homologues from the following bacteria are presented:

1. *Klebsiella pneumoniae* M6H 78578
2. *Salmonella paratyphi* A ATCC 9150
3. *Vibrio cholerae* serotype O1, Biotype El tor, Strain N16961
4. *Pseudomonas aeruginosa* PAO1
5. *Neisseria gonorrhoea* FA 1090
6. *Neisseria meningitidis* serogroup A Strain Z2491
7. *Streptococcus pyogenes* M1
8. *Bordetella pertussis* Tohama I
9. *Staphylococcus aureus* NCTC 8325
10. *Staphylococcus aureus* COL
11. *Porphyromonas gingivalis* W83
12. *Streptococcus mutans* UAB159
13. *Streptococcus pneumoniae* Type 4
14. *Clostridium difficile* 630 (epidemic type X)
15. *Camphylobacter jejuni* NCTC
16. *Bacillus anthracis* Ames
17. *Mycobacterium avium* 104
18. *Corynebacterium diphtheriae*
19. *Chlamydia trachomatis* MoPn

Klebsiella pneumoniae M6H 78578 (119 aa)
Amino acid sequence:
VVKLAFPRELRLLTFSHFTFVFQQPQRAGTPQITILGRLNSLGHPRIGLTVAKKNVKRAHERNRIKRLTRESFRLRQ
HELPPMDFVVVAKRGVADLDNRALSEALEKLWRRHCRLARGS
Nucleotide sequence (plus strand):
CGTCGTCGTGCTAAAGGCCGCGCTCGTCTGACCGTTTCCAAGTAATAAAGCTAACCCTGCGTGGTTAAGCTCGCATT
TCCCAGGGAGTTACGCTTGTTAACTCCCAGTCATTTCACTTTCGTCTTCCAGCAGCCACAACGGGCTGGCACGCCGC
AAATCACCATCCTCGGCCGCCTGAATTCGCTGGGGCATCCCCGCATCGGTCTCACCGTCGCCAAGAAAAACGTGAAA
CGCGCACATGAACGCAATCGGATTAAACGTCTGACGCGTGAAAGTTTTCGTTTGCGTCAACATGAACTCCCGCCAAT
GGATTTCGTGGTGGTGGCGAAAAGAGGGGTTGCCGACCTCGATAACCGTGCTCTCTCGGAAGCGTTGGAAAAATTAT
GGCGCCGCCATTGTCGCCTGGCTCGCGGGTCCTGATCGGCCTGATTCGAGTTTATCAGCGCCTGATTAGTCCGCTAC
TCGGGCCGCATTGTC
Sequence origin: Washington University; Contig 632

Salmonella paratyphi A ATCC 9150 (110 aa)
Amino acid sequence:
VTFVNSRSPHIRLPATSTGCTPQITILGRLNSLGHPRIGLTVAKKNVRRAHERNRIKRLTRESFRLRQHELPAMDFV
VVAKKGVADLDNRALSEALEKLWRRHCRLARGS
Nucleotide sequence (plus strand):
CTGACCGTTTCCAAGTAATAAAGCTAACCCCTGAGTGGTTAAGCTCGCATTTCCCAGGGAGTTACGTTTGTTAACTC
CCGCTCATTTCACATTCGTCTTCCAGCAACCTCAACGGGCTGCACGCCGCAAATCACCATCCTCGGCCGCCTGAATT
CGCTGGGGCATCCCCGTATCGGTCTTACCGTCGCCAAGAAAAATGTTCGACGTGCGCATGAACGCAACCGGATTAAA
CGTCTGACGCGTGAAAGCTTCCGTCTGCGCCAGCATGAACTTCCTGCAATGGATTTCGTGGTGGTGGCGAAAAAAGG
GGTTGCCGACCTCGATAACCGTGCTCTCTCGGAAGCGTTGGAAAAATTATGGCGCCGCCACTGTCGCCTGGCTCGCG
GGTCCTGATAGCCCTTATTCGGGTCTATCAACGCCTGATCAGTCCGCTGCTTGGGCCGCATTGTCGTTTC
Sequence origin: Washington University;

Vibrio cholerae serotype O1, Biotype El Tor, Strain N16961 (122 aa)
Amino acid sequence:
SRIILSTYAFNRELRLLTPEHYQKVFQQAHSAGSPHLTILARANNLSHPRLGLAVPKKQIKTAVGRNRFKRICRESF
RLHQNQLANKDFVVIAKKSAQDLSNEELFNLLGKLWQRLSRPSRG
Nucleotide sequence (minus strand): *NO INITIATOR CODON BEFORE STOP*
GGCAGCGTGGGCCGATAAGTGGACTAATAAACCACTGGTAAAGTTTTACAATACCAATGGCTAACCACGAGAAGGGC
GAGAGAGGCGTTGCCATAGTTTGCCAAGCAAGTTAAACAGTTCTTCATTGCTCAAATCTTGCGCGCTCTTTTTGGCG
ATGACAACAAAATCTTTGTTAGCCAGTTGATTTTGATGTAAGCGAAAGCTTTCTCTGCAAATACGTTTGAATCGATT
ACGGCCGACGGCAGTTTTGATCTGCTTTTTAGGAACCGCGAGTCCCAAACGAGGATGAGAAAGGTTATTAGCGCGAG
CGATGATTGTGAGATGAGGAGAACCAGCACTGTGAGCTTGCTGGAAGACTTTTTGATAATGTTCGGGAGTTAACAAA
CGTAACTCCCGATTGAATGCGTACGTACTCAAAATAATTCGAGATTATTTTGACAGGCGCTTACGGCCTTTTGCACG
ACGTGCATTCAGAACTTTACGACCGTTCGC
Sequence origin: TIGR

Pseudomonas aeruginosa PAO1 (135 aa)
Amino acid sequence:
VVSRDFDRDKRLLTARQPSAVFDSPTGKVPGKHVLLLARENGLDHPRLGLVIGKKNVKLAVQRNRLKRLIRESPRHN
QETLAGWDIVVIARKGLGELENPELHQQFGKLWKRLLRNRPRTESPADAPGVADGTHA
Nucleotide sequence (plus strand):
TCTGTCGCGTCGTCGCGCCAAAGGCCGTAAGCGTCTGACCGTCTGATTTATCCGGTACGGGTGGTGAGTCGGGACTT
CGACCGGGACAAGCGTCTACTGACAGCCCGGCAATTCAGCGCAGTCTTCGACTCTCCGACCGGCAAGGTCCCCGGCA
AGCACGTCCTGCTGCTGGCGCGCGAGAACGGTCTCGATCACCCCCGCCTGGGCCTGGTGATCGGCAAGAAGAACGTC
AAGCTCGCCGTCCAGCGCAATCGCCTCAAACGCCTGATCCGCGAATCGTTCCGCCATAACCAGGAAACCCTGGCTGG
CTGGGATATCGTGGTGATCGCGCGCAAAGGCCTGGGCGAACTGGAAAATCCGGAGCTGCACCAGCAGTTCGGCAAGC
TCTGGAAACGCCTGTTGCGCAATCGACCTCGCACGGAAAGCCCTGCTGACGCCCCTGGCGTGGCCGACGGTACTCAT
GCATAGGTCGATGCCCGCGCATCCCGATCCCTGTAGTGTCATCCCCCCTTCGATGACCCGGCACCG
Sequence origin: Pathogenesis & University of Washington; Contig 54

Neisseria gonorrhoea PA 1090 (123 aa)
Amino acid sequence:
VILDYRFGRQYRLLKTDDFSSVFAFRNRRSRDLLQVSRSNGNGLDHPRIGLVVGKKTAKRANERNYMKRVIRDWFRL
NKNRLPPQDFVVRVRRKFDRATAKQARAELAQLMFGNPATGCGKQV
Nucleotide sequence (minus strand):
ATGTTCCTTGTATGGGAAACCCGTTGCCGTCTGAACCTTGCCTGCAGGGTACCGTTCTGATCATACCTGTTTCCCGC
ATCCGGTTGCGGGGTTGCCGAACATGAGTTGTGCCAGTTCCGCCCTTGCCTGTTTTGCGGTAGCCCTGTCGAATTTC
CGGCGGACGCGCACGACGAAATCCTGAGGCGGCAGCCGGTTTTTGTTCAATCTGAACCAGTCGCGGATGACGCGTTT
CATATAGTTCCGCTCGTTGGCGCGTTTGGCGGTTTTTTTGCCGACCACCAGACCGATGCGGGGATGGTCCAGCCCGT
TGCCGTTTGAGCGCGAAACTTGCAGCAGGTCGCGGCTGCGGCGGTTTCTGAATGCAAAAACGGATGAAAAATCATCC
GTTTTTAACAAGCGGTACTGCCTTCCGAAGCGGTAGTCCAAAATTACACTGCCAGGCGTTTGCGGCCTTTGGCACGG
CGTGCGGCCAATACTGCGCGTCCGCCGCGT
Sequence origin: University of Oklahoma ACGT; Contig 60

Neisseria meningitidis serogroup A Strain Z2491 (123 aa)
Amino acid sequence:
VILDYRFGRQYRLLKTDDFSSVFAFRNRRSRDLLQVSRSNGNGLDHPRIGLVVGKKTAKRANERNYMKRVIRDWFRL
NKNRLPPQDFVVRVRRKFDRATAKQARAELAQLMFGNPATGCRKQA
Nucleotide sequence (minus strand):
TGTTCCTTAGTATGGGAAACCCGTTGCCGTCTGAACCTTGCCTGCAGAGTACCGTTCTGATCATGCCTGTTTCCTGC
ATCCGGTTGCGGGGTTGCCGAACATGAGTTGTGCCAGTTCCGCCCTTGCCTGTTTTGCGGTAGCCCTGTCGAATTTA
CGGCGGACGCGCACGACGAAATCCTGCGGCGGCAGCCGGTTTTTGTTCAATCTGAACCAGTCGCGGATGACGCGCTT
CATATAATTTCGTTCGTTGGCGCGTTTGGCGGTTTTTTTGCCGACCACCAGACCGATGCGGGGATGATCCAGCCCGT
TGCCGTTTGAACGCGAAACTTGCAGCAGGTCGCGGCTGCGGCGGTTTCTGAATGCAAAAACGGATGAAAAATCATCC
GTTTTCAACAAGCGGTACTGCCTTCCGAAGCGGTAGTCCAAAATTACACCGCCAGGCGTTTGCGGCCTTTGGCGCGC
CGTGCGGCCAATACTGCGCGTCCGCCGCGC
Sequence origin: Sanger centre & Oxford University

Streptococcus pyogenes M1 (113 aa)
Amino acid sequence:
VKREKDFQAIFKDGKSTANRKFVIYHLNRGQDHFRVGISVGKKIGNAVTRNAVKRKIRHVIMALGHQLKSEDFVVIA
RKGVESLEYQELQQNLEHVLKLAQLLEKGFESEEKH
Nucleotide sequence (minus strand):
GTTACCTCACCACGACCACAGGCCACTAATAATAGAACTAAGGGGACTATTCTTGCAATTTTAATGTTTTTCTTCAC
TCTCAAAACCTTTCTCAAGCAATTGTGCTAACTTTAAAACATGATGTAAATTTTGTTGAAGCTCTTGATACTCCAAA
GATTCGACACCCTTACGGGCAATCACCACGAAATCCTCTGACTTCAGCTGATGCCCTAATGCCATGATAACATGACG
TATCTTTCGTTTGACTGCATTTCTGGTGACTGCATTTCCTATTTTTTTACCGACAGAAATACCCACACGGAAGTGGT
CTTGGCCTCTATTTAAATGATAAATGACAAATTTTCGATTTGCTGTACTTTTTCCATCCTTAAATATGGCTTGGAAA
TCTTTCTCACGCTTGACACGATAGGTCTTCTTCAAAATTTAACTCCAATATCTAAATTATTACCATTATACCACATC
Sequence origin: University of Oklahoma ACGT; Contig 7

Bordetella pertussis Tohama I (123 aa)
Amino acid sequence:
MPRATLPAEARLHRPSEFAAALKGRRLARGAFFIVSASPCAPADDQPARARLGLVIAKRFAARAVTRNTLKRVIREA
FRARRLALPAQDYVVRLHSKLTPASLTALKRSARAEVDAHFTRIAR
Nucleotide sequence (minus strand):
CCACCCAGGGGCTGAGGAAGTACCGGTAAAACCGGATCGGGGCGATAAGCAGTCTCCTGATCATCGCGCTATCCGTG
TGAAGTGAGCATCTACTTCGGCGCGCGCCGAGCGTTTCAGGGCCGTGAGGCTTGCCGGTGTCAGCTTGCTGTGCAGC
CGCACCACGTAATCCTGGGCCGGCAGGGCAAGCCGGCGAGCCCGGAACGCTTCGCGGATGACCCGCTTCAAGGTATT
GCGCGTCACGGCGCGGGCGGCAAAACGCTTGGCGATCACCAGGCCCAGGCGCGCGCGCGCCGGCTGGTCATCAGCAG
GGGCACAGGGCGAGGCGCTGACAATAAAGAAAGCCCCTCGGGCCAGTCGCCGGCCTTTGAGGGCGGCGGCAAACTCG
GAGGGGCGATGCAATCGCGCCTCCGCAGGGAGCGTGGCGCGCGGCATCGGTGACGTGACGGAGACTGGCGACGGGGC
CGGCGGCGATGCTCCTGTTACAGGCAATCC
Sequence origin: Sanger centre & MDS; Contig 267

Staphylococcus aureus NCTC 8325 (117 aa)
Amino acid sequence:
MLLEKAYRIKKNADFQRIYKKGHSVANRQFVVYTCNNKEIDHFRLGISVSKKLGNAVLRNKIKRAIRENFKVHKSHI
LAKDIIVIARQPAKDMTTLQIQNSLEHVLKIAKVFNKKIK
Nucleotide sequence (plus strand):
GTTATAAGCTCAATAGAAGTTTAAATATAGCTTCAAATAAAAACGATAAATAAGCGAGTGATGTTATTGGAAAAAGC
TTACCGAATTAAAAAGAATGCAGATTTTCAGAGAATATATAAAAAAGGTCATTCTGTAGCCAACAGACAATTTGTTG
TATACACTTGTAATAATAAAGAAATAGACCATTTTCGCTTAGGTATTAGTGTTTCTAAAAAACTAGGTAATGCAGTG
TTAAGAAACAAGATTAAAAGAGCAATACGTGAAAATTTCAAAGTACATAAGTCGCATATATTGGCCAAAGATATTAT
TGTAATAGCAAGACAGCCAGCTAAAGATATGACGACTTTACAAATACAGAATAGTCTTGAGCACGTACTTAAAATTG
CCAAAGTTTTTAATAAAAAGATTAAGTAAGGATAGGGTAGGGGAAGGAAAACATTAACCACTCAACACATCCCGAAG
TCTTACCTCAGA
Sequence origin: University of Oklahoma ACGT; Contig 561

Staphylococcus aureus COL (117 aa)
Amino acid sequence:
MLLEKAYRIKKNADFQRIYKKGHSVANRQFVVYTCNNKEIDHFRLGISVSKKLGNAVLRNKIKRAIRENFKVHKSHI
LAKDIIVIARQPAKDMTTLQIQNSLEHVLKIAKVFNKKIK
Nucleotide sequence (plus strand):
GTTATAAGCTCAATAGAAGTTTAAATATAGCTTCAAATAAAAACGATAAATAAGCGAGTGATGTTATTGGAAAAAGC
TTACCGAATTAAAAAGAATGCAGATTTTCAGAGAATATATAAAAAAGGTCATTCTGTAGCCAACAGACAATTTGTTG
TATACACTTGTAATAATAAAGAAATAGACCATTTTCGCTTAGGTATTAGTGTTTCTAAAAAACTAGGTAATGCAGTG
TTAAGAAACAAGATTAAAAGAGCAATACGTGAAAATTTCAAAGTACATAAGTCGCATATATTGGCCAAAGATATTAT
TGTAATAGCAAGACAGCCAGCTAAAGATATGACGACTTTACAAATACAGAATAGTCTTGAGCACGTACTTAAAATTG
CCAAAGTTTTTAATAAAAAGATTAAGTAAGGATAGGGTAGGGGAAGGAAAACATTAACCACTCAACACATCCCGAAG
TCTTACCTCAGA
Sequence origin: TIGR;

Porphyromonas gingivalis W83 (137 aa)
Amino acid sequence:
MTSPPTFGLSKSERLYLRDEINTVFGEGKAFVVYPLRVVYRLGSEHRVAYSSMLVSVAKKRFRRAVKRNRVKRLVRE
AYRLNKHLLNDVLQERQIYATIAFMVVSDELPDFRTVERAMQKSLIRIAGNVPSSALKNE
Nucleotide sequence (minus strand):
AGAAGAAAATGGGGAGCAGTAAGAGTTGCACGAGAAAAGCCTTGATCAGTCGCATCGTATTTACTCGTTTTTCAAAG
CCGATGAAGGTACATTTCCGGCAATTCTGATCAGACTCTTTTGCATCGCTCTCTCCACTGTACGAAAGTCAGGAAGT
TCATCCGATACTACCATAAATGCAATAGTAGCATAGATCTGTCTCTCTTGGAGGACATCGTTCAGGAGGTGTTTGTT
GAGCCGATAAGCCTCCCTGACCAAACGCTTGACCCTATTGCGCTTCACGGCTCGCCTAAACCTTTTCTTTGCTACGC
TTACCAGCATGGAGGAATATGCAACTCGATGCTCCGATCCCAGACGGTAGACTACGCGTAGAGGATAAACGACAAAC
GCCTTGCCTTCGCCAAAGACCGTATTGATTTCATCGCGAAGATAGAGGCGTTCGCTTTTGGATAGGCCGAATGTAGG
CGGAGAGGTCATTTCCCGTTGAGGTAATCCTCTAATGCCATAGCCATAGAAGGATATTGCTCGGTCGGCGCA
Sequence origin: TIGR & Forsyth Dental Center

Streptococcus mutans UAB159 (119 aa)
Amino acid sequence:
VLKKAYRVKSDKDFQAIFTEGRSVANRKFVVYSLEKDQSHYRVGLSVGKRLGNAVVRNAIKRKLRHVLMELGPYLGT
QDFVVIARKGVEELDYSTMKKNLVHVLKLAKLYQEGSIREKE
Nucleotide sequence (plus strand):
AGATTTTTGGCTTTTTCTCATTTTATGATATAATAGTGATAATTTAAATATTGGAGTCATGTTTTGAAAAAAGCCTA
TCGCGTTAAAAGTGATAAAGATTTTCAGGCAATTTTTACTGAAGGACGAAGTGTTGCCAATCGGAAATTTGTTGTCT
ATAGTTTAGAAAAAGATCAAAGTCACTATCGTGTTGGACTTTCAGTTGGAAAAAGATTAGGAAATGCTGTCGTTAGA
AATGCGATTAAACGAAAATTGCGCCATGTCCTTATGGAACTTGGTCCTTATTTAGGCACTCAAGATTTTGTTGTTAT
TGCTAGAAAAGGTGTTGAGGAACTTGATTATAGCACGATGAAAAAAAATCTGGTTCATGTTTTAAAACTGGCTAAAC
TGTATCAGGAAGGATCTATTCGTGAAAAAGAA
Sequence origin: University of Oklahoma ACGT; Contig 299

Streptococcus pneumoniae Type 4 (124 aa)
Amino acid sequence:
VLKKNFRVKREKDFKAIFKEGTSFANRKFVVYQLENQKNRFRVGLSVSKKLGNAVTRNQIKRRIREHIIQNAKGSLVE
DVDFVVIARKGVETLGYAEMEKNLLHVLKLSKIYREGNGSEKETKVD
Nucleotide sequence (minus strand):
TCGCTAGTTACCCCATTAGTCGCACAGGCTGTCATGATTAACAGAGACAGTCCTAGCAAACTAGTCAACTTTAGTTT
CTTTTTCACTCCCATTTCCTTCCCGGTAAATCTTTGATAATTTTAATACATGGAGTAGATTTTTCTCCATCTCTGCG
TATCCCAAGGTTTCGACTCCTTTTCGAGCAATGACAACAAAGTCGACATCTTCTACCAGACTCCCTTTTGCATTCTG
GATAATATGCCGAATCCGTCGCTTAATTTGATTTCTAGTGACGGCATTCCCCAGTTTTTTGCTAACTGATAGACCTA
CTCGAAAACGGTTTTTCTGGTTTTCTAATTGGTAGACCACAAATTTGCGATTAGCAAAACTTGTCCCCTCCTTGAAA
ATCGCCTTAAAATCTTTCTCTCTTTTTACACGAAAGTTTTTCTTCAAAACTCAACTCCATCTATTAAATTACTACTA
TTATACCATATTTTTCAAAAAAGCCAATCATAG
Sequence origin: TIGR;

Clostridium difficile 630 (epidemic type X) (114 aa)
Amino acid sequence:
MDFNRTKGLKKDSDFRKVYKHGKSFANKYLVIYILKNKSDYSRVGISVSKKVGKAITRNRVRRLIKEAYRLNIDEKI
KPGYDIVFIARVSSKDATFKDIDKSIKNLVKRTDISI
Nucleotide sequence (minus strand):
TCCTTTAATATATAAATTATTTTATTCAAAGTCATTAACCTCCATATTTATAGCATACAATTAAATAGAAATATCCG
TTCTTTTAACTAAATTTTTTATAGACTTGTCTATGTCTTTAAAAGTAGCATCCTTACTAGATACCCTTGCTATAAAT
ACTATATCATATCCAGGCTTAATTTTTTCATCAATATTTAATCTGTAGGCTTCTTTTATTAATCTTCTTACTCTATT
CCTAGTAATAGCTTTTCCTACTTTTTTTGAAACAGAAATACCTACTCTACTATAATCTGATTTATTTTTAAGTATAT
ATATTACTAAATATTTGTTTGCAAAAGATTTGCCGTGTTTATATACTTTTCTAAAATCAGAGTCTTTTTTCAACCCT
TTAGTCCTATTAAAGTCCATAGTTAACCTCCATAAACACAGCTATGAATCGTAATTATTTACACAAAAAGGCCACCT
TTG
Sequence origin: Sanger centre; Contig 975

Campylobacter jejuni NCTC (108 aa)
Amino acid sequence:
VKNFDKFSTNEEFSSVYKVGKKWHCEGVIIFYLNSYEKKIAVVASKKVGKAVVRNRSKRILRALFAKFERYLQDGKY
IFVAKNEITELSFSRLEKNLKWGLKKLECFK
Nucleotide sequence (minus strand):
AAGCAGCGGGTTTTAAAGGGCTTAAGAATTTCTGATAAAAACGGAGTATTTTTAGGCATATCATTTGAAACATTCTA
GTTTTTTCAATCCCCATTTTAGATTTTTTTCTAACCTAGAAAAAGAAAGTTCAGTGATTTCATTTTTAGCTACAAAA
ATATATTTGCCATCTTGAAGATATCTTTCAAACTTAGCAAACAAAGCTCTTAAAATTCGTTTTGAACGATTTCTAAC
CACTGCTTTTCCAACTTTTTTACTAGCAACAACTGCTATTTTTTTTTCATAACTATTCAGATAAAAAATGATCACAC
CTTCGCAATGCCATTTTTTGCCTACTTTATATACAGATGAAAATTCCTCGTTTGTGCTAAATTTATCAAAATTTTTC
ACACAGCAAGTCTTTTTCTACCTTTAGCGCGTCTTGCATTGATCACTTTGCGACCATTTTTA
Sequence origin: Sanger centre & MDS

Bacillus anthracis Ames (119 aa)
Amino acid sequence:
MKKKHRIKKNDEFQTVFQKGKSNANRQFVVYQLDKEEQPNFRIGLSVSKKIGNAVVRNRIKRMIRQSITELKDEIDS
GKDFVIIARKPCAEMTYEELKKSLIHVFKRSGMKRIKSSVRK
Nucleotide sequence (minus strand):
TAAACCTAATTTCTTTTTCAAAGCCTACTCCTCCTTGTATCGGTATGTATATAGTGTAATTCATTTCCTTACGCTAC
TTTTTATTCTTTTCATACCAGAGCGTTTAAAGACATGAATTAAGCTTTTCTTTAATTCTTCATATGTCATCTCTGCA
CAAGGCTTCCTTGCTATTATAACAAAATCTTTTCCAGAATCTATCTCATCTTTTAATTCTGTGATCGACTGGCGAAT
CATACGTTTAATTCGGTTACGCACTACTGCATTTCCTATCTTCTTGCTGACAGAAAGGCCAATACGAAAGTTTGGCT
GCTCTTCTTTATCTAGTTGATAGACAACAAATTGACGATTCGCATTCGATTTTCCTTTTTGAAAAACCGTCTGGAAT
TCATCATTCTTTTTTATACGATGTTTTTTCTTCATATCAATTGACACTCCTGTAGTTCATCAGCGGAAATTCACTAT
TATTAGAAAAAAAGACCA
Sequence origin: TIGR;

Mycobacterium avium 104 (119 aa)
Amino acid sequence:
VLPARNRMTRSTEFDATVKHGTRMAQPDIVVHLRRDSEPDDESAGPRVGLVVGKAVGTAVQRHRVARRLRHVARALL
GELEPSDRLVIRALPGSRTASSARLAQELQRCLRRMPAGTGP
Nucleotide sequence (minus strand):
GTCCGCGGGCGACGGTTCGGCCGGCGCCGCGAATGGCCGCGCCCGACCGCGCCGGTCCGGTCACGGCCCGGTTCCCG
CCGGCATGCGCCGCAGGCACCGCTGCAGTTCCTGCGCCAGGCGCGCCGACGACGCGGTCCGGCTTCCGGGCAGCGCG
CGAATCACCAGCCGGTCGGATGGTTCGAGTTCGCCGAGCAGGGCCCGGGCCACGTGACGCAGCCGGCGGGCCACGCG
GTGTCGTTGCACCGCCGTCCCGACGGCCTTCCCGACGACCAGCCCGACCCGTGGGCCCGCCGGATTCGTCGTCGGGTT
CGGAGTCGCGCCGGAGGTGGACGACGATGTCGGGCTGCGCCATGCGGGTTCCGTGCTTCACCGTCGCGTCAAACTCG
GTTGACCGCGTCATGCGGTTGCGTGCGGGAAGCACCGCGAAAGACCTGACGTGCGATCAGGCAGAGAGCGCGCGGCG
ACCCTTGCGGCGCCGACC
Sequence origin: TIGR;

Corynebacterium diphtheriae (129 aa)
Amino acid sequence:
VTLTSSNRTTVLPSQHKLSNSEQFRATIRKGKRAGRSTVVLHFYAEATAGNLATAGGPRFGLVVSKAVGNAVTRHRV
SRQLRHVVIAMKDQFPASSHVVVRAIPPAATASYEELRADVQAALDKLNRKR
Nucleotide sequence (plus strand):
CCGGTCGCGCAATCGTGGCTGCACGTCGTAACAAGGGTCGTAAGAGCCTGACCGCTTAAGGTCACTCTTACAAGCTC
GAATAGAACGACGGTGCTACCTTCACAGCACAAGCTCAGCAATTCCGAACAGTTCCGCGCAACGATTCGGAAGGGCA
AGCGTGCTGGGAGGAGCACCGTCGTTCTTCATTTTTATGCTGAGGCGACCGCGGGCAACCTTGCAACCGCAGGCGGC
CCGCGATTCGGCCTCGTTGTGTCCAAGGCTGTTGGAAATGCTGTGACTCGTCACCGTGTTTCGCGGCAGTTAAGGCA
CGTAGTAATCGCTATGAAAGACCAGTTCCCAGCGTCATCCCATGTTGTTGTGAGGGCGATACCGCCAGCGGCGACAG
CAAGTTATGAGGAGTTGCGGGCAGATGTGCAGGCAGCACTCGACAAGCTCAACCGCAAGCGATAAGGCGGTTACTCG
CCCTCGTGGGCTGGTTAGTCGCGCATTGTTTGATGCGGTGCGGTTCTA
Sequence origin: Sanger centre; Contig 390

Chlamydia trachomatis MoPn (119 aa)
Amino acid sequence:
VHRLTLPKSARLLKRKQFVYVQRCGQYCRTDQATLRIVPSRHSNIRKVGVTVSKKFGKAHQRNRFKRIVREAFRHVR
PNLPACQVVVSPKGGTLPNFGKLSADLLKHIPEALPLVTSSK
Nucleotide sequence (plus strand):
GCTACAAAAAGTGGAAGAAATCTTTTAAATCGTCGTCGCCGTCACGGCAGACATTCCTTAATTGATCTCTAAGATCT
TTCATTTGTGCATCGGTTAACTCTACCTAAAAGTGCCCGCCTATTGAAACGTAAACAATTTGTTTACGTGCAGCGTT
GTGGGCAATATTGTCGTACTGATCAGGCAACTTTACGAATAGTTCCTTCTCGTCATTCGAACATCCGTAAAGTAGGG
GTTACTGTTTCTAAAAAATTTGGGAAAGCCCATCAGCGCAATCGCTTTAAAAGAATTGTGCGAGAGGCTTTTAGGCA
TGTGCGACCAAATCTTCCCGCATGTCAAGTGGTAGTGTCTCCTAAAGGGGGCACTCTACCAAATTTTGGTAAACTAT
CCGCGGATCTTCTTAAGCATATTCCAGAGGCTTTGCGTGTCGTTACTTCTTCTAAGTAGTTTTTTATTTTGGTCAAA
AAATAAAAAACCATTCCACGCTATAGAGGCATGGAATGGGAA
Sequence origin: TIGR & Manitoba University;

New Sequences:

Pasteurella multocida PM70 (130? aa)
Amino acid sequence:
?IVISIILIPIGVIKLNPSRELRLLTPLHFKYVFEQPFRASTPELTILARPNNLAHPRLGLTVAKKHLKKAHDRNRI
KRLCRESFRLAQYKLPNCDFVIVAKQGIGKLDNRTLTQTLDKLWQRHIRLAQKS
Nucleotide sequence (plus strand):
ATCGTTATCAGCATAATCTTAATCCCTATTGGTGTGATTAAGCTGAATTTTTCGAGGGAGTTACGTTTGTTAACTCC
CCTTCATTTTAAATACGTCTTCGAACAGCCGTTCCGTGCTAGTACACCTGAACTTACCATTCTTGCTCGTCCCAATA
ATCTCGCTCATCCTCGCTTAGGGTTAACTGTCGCGAAAAAGCATTTAAAAAAAGCACATGATCGCAATCGCATCAAA
CGCTTATGCCGAGAAAGTTTCCGCCTAGCACAGTATAAACTCCCCAATTGCGATTTTGTTATTGTGGCGAAACAGGG
AATTGGTAAATTAGACAACAGGACACTCACACAAACATTGGATAAATTATGGCAAAGACACATTCGCTTAGCTCAAA
AATCTTGA

Haemophilus ducreyi strain 35000HP (130? aa)
Amino acid sequence:
?SVSKVTSVNKLTFSRELRLLAPIQFKAVFEQPYRASTAELTILARQNCVNTPRLGLTVAKKHLKRAHDRNRIKRIV
RESFRLKQHQLPNFDFVFVAKHGIGKLDNATLFATIDKLWTRHIRLSQQAQSKN
Nucleotide sequence (minus strand):
TTAATTTTTGCTTTGTGCTTGTTGACTGAGGCGAATATGACGAGTCCATAATTTATCTATGGTTGCGAAAAGCGTAG
CATTATCTAGTTTACCAATCCCATGCTTGGCAACAAAGACAAAGTCAAAATTAGGTAATTGATGTTGTTTTAAACGG
AAGCTTTCCCGCACAATACGTTTGATCCGATTGCGATCGTGAGCACGTTTTAAATGCTTTTTAGCAACGGTTAACCC
AAGACGAGGCGTATTAACGCAATTTTGACGAGCAAGAATAGTAAGTTCAGCTGTGCTAGCACGATATGGTTGTTCAA
ACACGGCTTTGAATTGAATGGGAGCTAACAAACGTAGCTCCCGAGAAAACGTTAGCTTATTCACTGACGTGACTTTG
CTGACACT

Chlamydia muridarum (? aa)
Amino acid sequence:
RLTLPKSARLLKRKQFVYVQRCGQYCRTDQATLRIVPSRHSNIRKVGVTVSKKPGKAHQRNRFKRIVREAFRHVRPN
LPACQVVVSPKGG
Nucleotide sequence (? strand):

Exhibit F

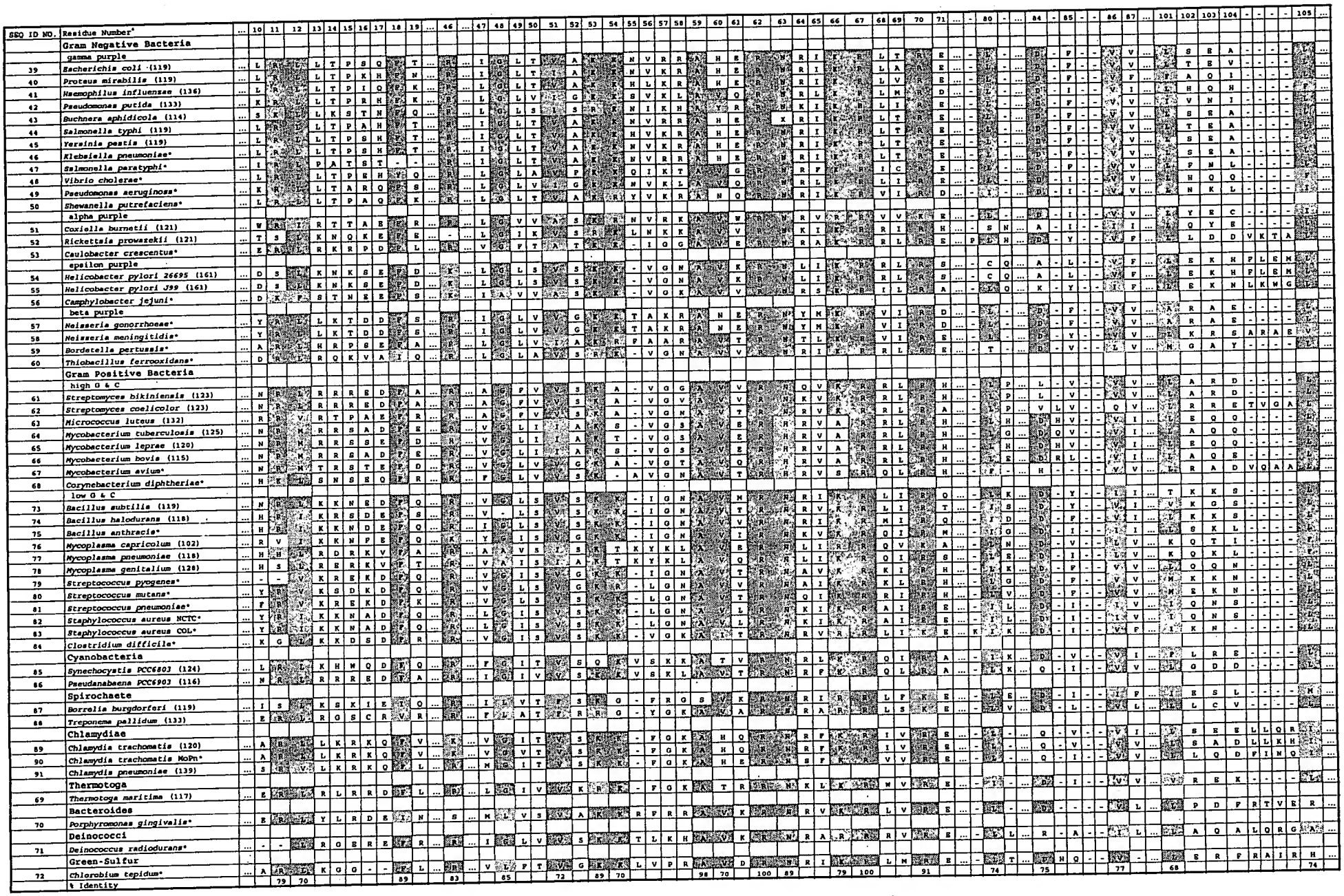


FIG. 1

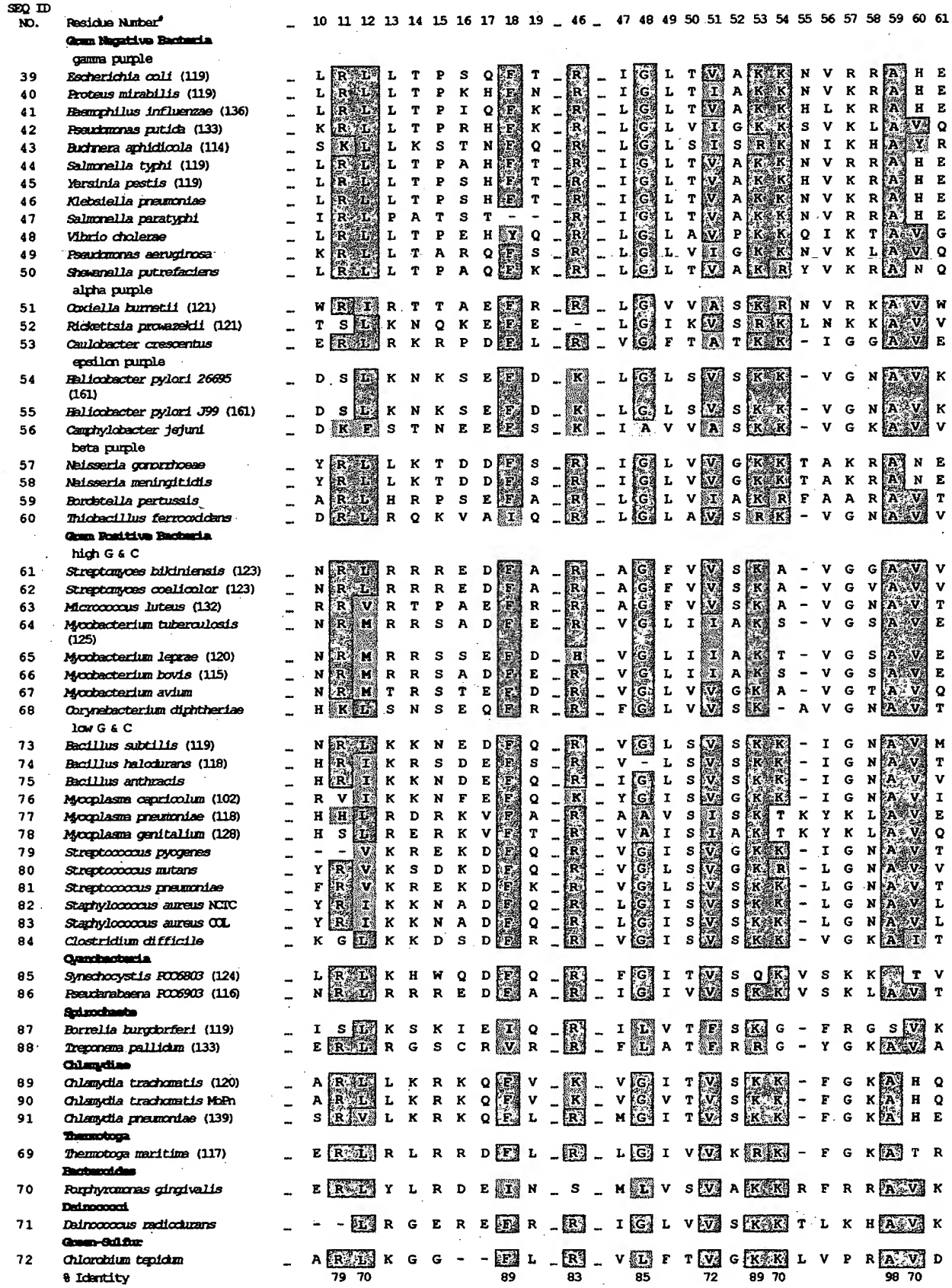

SEQ ID NO.
Residue Number
10 11 12 13 14 15 16 17 18 19 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61
Gram Negative Bacteria
gamma purple
39 Escherichia coli (119)
40 Proteus mirabilis (119)
41 Haemophilus influenzae (136)
42 Pseudomonas putida (133)
43 Buchnera aphidicola (114)
44 Salmonella typhi (119)
45 Yersinia pestis (119)
46 Klebsiella pneumoniae
47 Salmonella paratyphi
48 Vibrio cholerae
49 Pseudomonas aeruginosa
50 Shewanella putrefaciens
alpha purple
51 Coxiella burnetii (121)
52 Rickettsia prowazekii (121)
53 Caulobacter crescentus
epsilon purple
54 Helicobacter pylori 26695 (161)
55 Helicobacter pylori J99 (161)
56 Camphylobacter jejuni
beta purple
57 Neisseria gonorrhoeae
58 Neisseria meningitidis
59 Bordetella pertussis
60 Thiobacillus ferrooxidans
Gram Positive Bacteria
high G & C
61 Streptomyces bikiniensis (123)
62 Streptomyces coelicolor (123)
63 Micrococcus luteus (132)
64 Mycobacterium tuberculosis (125)
65 Mycobacterium leprae (120)
66 Mycobacterium bovis (115)
67 Mycobacterium avium
68 Corynebacterium diphtheriae
low G & C
73 Bacillus subtilis (119)
74 Bacillus halodurans (118)
75 Bacillus anthracis
76 Mycoplasma capricolum (102)
77 Mycoplasma pneumoniae (118)
78 Mycoplasma genitalium (128)
79 Streptococcus pyogenes
80 Streptococcus mutans
81 Streptococcus pneumoniae
82 Staphylococcus aureus NCTC
83 Staphylococcus aureus COL
84 Clostridium difficile
Cyanobacteria
85 Synechocystis PCC6803 (124)
86 Pseudanabaena PCC6903 (116)
Spirochaete
87 Borrelia burgdorferi (119)
88 Treponema pallidum (133)
Chlamydiae
89 Chlamydia trachomatis (120)
90 Chlamydia trachomatis MoPn
91 Chlamydia pneumoniae (139)
Thermotoga
69 Thermotoga maritima (117)
Bacteroides
70 Porphyromonas gingivalis
Deinococci
71 Deinococcus radiodurans
Green-Sulfur
72 Chlorobium tepidum
% Identity
79 70
89
83
85
72
89 70
98 70

FIG. 1 Continued